UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

IMPROVING THE PREDICTION OF NOCTURNAL CONVECTION THROUGH THE ASSIMILATION OF NOVEL DATASETS: OBSERVATION IMPACTS AND ERROR TREATMENT

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

Samuel Kyle Degelia
Norman, Oklahoma
2021

IMPROVING THE PREDICTION OF NOCTURNAL CONVECTION THROUGH THE ASSIMILATION OF NOVEL DATASETS: OBSERVATION IMPACTS AND ERROR TREATMENT

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Xuguang Wang, Chair
Dr. Prakash Vedula
Dr. Michael Richman
Dr. Alan Shapiro
Dr. David Stensrud

Acknowledgments

First, I would like to thank my advisor, Dr. Xuguang Wang. Her expertise has opened the door for countless opportunities throughout my last six years as a graduate student. Without her aid, I would not be anywhere close to the scientist I am now. I also want to thank Dr. David Stensrud who served as co-chair of my M.S. committee and continues to act as a mentor despite his countless responsibilities. I would also like to acknowledge the other members of my committee - Drs. Prakash Vedula, Alan Shapiro, and Michael Richman. Next, I thank Drs. David Turner, and Petra Klein for providing additional guidance throughout my time at the University of Oklahoma. A special mention should also go to Dr. Aaron Johnson who served as an additional mentor and tech support guru during my first year as a graduate student. Additional thanks also go to the support desk at the Computational and Information Systems Laboratory (CISL) for responding to my barrage of questions.

Next, I also want to thank past and present members of the Multi-scale data Assimilation and Predictability (MAP) laboratory at OU including Bo Huang, Xu Lu, Yongming Wang, Hristo Chipilski, Junkyung Kay, Krishnamoorthy Chandramouli, and Nicholas Gasperoni. Our various chats, both scientific and non-, have proven very helpful in learning the ins and outs of graduate school.

On a personal note, I could never have made it this far without the support from my incredible wife, Gabrielle. Despite being thousands of miles away at times, she has never failed to brighten up even the toughest of days. Our wonderful dog Carly also deserves a special shoutout for her support and helpful distractions from work. Finally, my parents (Pat and Laura) have provided incredible support throughout my entire life, both financial and personal. I can only hope that this dissertation makes them proud.

Table of Contents

List of Tables

List of Figures

Abstract

Numerical weather prediction (NWP) models often fail to correctly forecast both the initiation and evolution of nocturnal convection. To improve our understanding of such events, researchers collected a unique dataset of thermodynamic and kinematic remote sensing profilers as part of the Plains Elevated Convection at Night (PECAN) experiment. This dissertation evaluates the impacts made to forecasts of nocturnal convection when assimilating this network that includes atmospheric emitted radiance interferometers (AERIs), Doppler lidars, radar wind profilers (RWPs), high-frequency rawinsondes, and mobile surface observations.

Using an advanced, ensemble-based data assimilation system, we first evaluate the impacts of these datasets for a single nocturnal convection initiation (CI) event. Compared to operational forecasts, assimilating the PECAN dataset improves the timing, location, and orientation of the CI forecast. We show that AERIs, RWPs, and rawinsondes produce the largest benefits by enhancing the moisture advection into the region of CI. The impacts of assimilating these datasets also remains large throughout the growth of the CI event into a mesoscale convective system (MCS). Assimilating Doppler lidar and surface data only slightly improves the CI forecasts by enhancing the convergence along an outflow boundary that partially forces the CI. While this case study suggests positive results from assimilating high-frequency profiling data, one single event cannot fully represent the wide diversity of mechanisms and environments that can lead to nocturnal convection.

To address additional types of nocturnal CI, we next expand our work into a systematic evaluation of the impact of assimilating a collocated network of high-frequency, ground-based thermodynamic and kinematic profilers collected during PECAN. For 13 nocturnal CI events, we find small but consistent improvements when assimilating thermodynamic observations collected by AERIs. Through midlevel cooling and moistening, assimilating the AERIs increases the skill for both nocturnal CI and precipitation forecasts. Assimilating composite kinematic datasets collected by Doppler lidars and RWPs instead results in slight degradations to the forecast quality, including decreases in skill and traditional contingency metrics. The impacts from assimilating thermodynamic and kinematic profilers often counteract each other, such that we find little impact on the detection of CI when both are assimilated. However, assimilating both datasets improves various properties of the CI events that are successfully detected (timing, distance, shape, etc.).

We hypothesize that a lack of flow-dependent methods to assign observation errors likely contributes to the forecast degradations for some cases. This theory motivates our final study which evaluates the impact of using novel methods for assigning observation errors when assimilating ground-based thermodynamic profilers. We find that a static error inflation method results in forecast degradations compared to a reference experiment where no remote sensing data are assimilated. These issues are partially resolved when adaptively inflating the observation errors or when using a method that computes the full observation error based on observation-space diagnostics. Flow-dependent extensions of each method are shown to further improve forecasts compared to their static counterpart by increasing observation errors for problematic retrievals. Assuming that the observation errors are correctly diagnosed, the results from this dissertation suggest that assimilating a network of ground-based thermodynamic profilers can greatly improve forecasts of nocturnal convection.

Chapter 1

Introduction

Multiple sections of this dissertation are direct excerpts of Degelia et al. (2019, ©American Meteorological Society) and Degelia et al. (2020, ©American Meteorological Society). This includes parts of Chapters 1, 2, 3, and 4.

1.1 Background and motivation

A nocturnal maximum in precipitation exists in the Great Plains of the United States (Surcel et al. 2010). Much of this rainfall occurs within mesoscale convective systems (MCS) that propagate eastward after developing over the higher terrain near the Rocky Mountains (Carbone et al. 2002; Trier et al. 2017). Locally-forced nocturnal convection initiation (CI) also contributes to the precipitation maximum, as Weckwerth and Romatschke (2019) find that CI events produce 30-60% of observed nocturnal rainfall in the Great Plains. Though this precipitation is beneficial for agricultural purposes, nocturnal convection also produces all types of severe weather (Horgan et al. 2007) with hail and wind being the most common threats (Reif and Bluestein 2017).

Despite the frequent occurrence of nocturnal convection, forecast skill for such events is typically lower than corresponding forecasts of daytime convection (Clark 2016; Keclik et al. 2017). Reif and Bluestein (2017) explain that many NWP systems are typically tuned specifically for mechanisms that initiate surface-based convection during the day, whereas nocturnal convection tends to be initiated by features above the boundary layer (Corfidi et al. 2008). As such, NWP models routinely fail to predict many mechanisms responsible for both the initiation and evolution of nocturnal convection (e.g., Davis et al. 2003; Johnson and Wang 2017; Johnson et al. 2017a; Stelten and Gallus 2017; Johnson et al. 2017b). For example, various studies find that NWP models often fail to correctly predict both the height and strength of the nocturnal low-level jet (LLJ; Storm et al. 2009; Shin and Hong 2011; Smith et al. 2015; Johnson and Wang 2017; Johnson et al. 2017a) which commonly contributes to the development of nocturnal convection through enhanced convergence at its terminus (Trier and Parsons 1993). Models also struggle to simulate elevated moist layers that provide a source of instability for nocturnal convection (Peters et al. 2017). Furthermore, many studies also document the inability of modern NWP systems to correctly predict atmospheric bores (Johnson et al. 2017b; Chipilski et al. 2020) that commonly aid in the maintenance of nocturnal MCSs (Haghi et al. 2018).

Assimilating surface and upper-air observations can resolve some issues in convective-scale NWP by improving analyses of both the pre-convective environment and the forcing mechanisms responsible for CI (e.g., Johnson et al. 2015; Degelia et al. 2018; Chen et al. 2020). Sobash and Stensrud (2015) find that assimilating mesonet observations for a daytime CI event can improve moisture errors and associated CI forecasts. However, few observations exist that frequently sample the 900-600 hPa layer where most nocturnal convection originates (Wilson and Roberts 2006; Weckwerth et al. 2019). Observations in this layer are typically provided by conventional rawinsondes which only collect data a few times per day, despite convective-scale environments evolving on the order of minutes to hours (Orlanski 1975).

To advance the prediction and understanding of nocturnal convection, scientists conducted the Plains Elevated Convection at Night (PECAN; Geerts et al. 2017) experiment in 2015. As a part of PECAN, various remote sensing profilers were deployed across the Great Plains, including Atmospheric Emitted Radiance Interferometers (AERIs; Feltz et al. 2003; Turner and Löhnert 2014), radar wind profilers (RWPs; Ecklund et al. 1988), and Doppler lidars (Menzies and Hardesty 1989). Although their errors are often larger than corresponding in-situ measurements, remote sensing instruments can provide thermodynamic and kinematic data throughout much of the lower troposphere multiple times per hour. As such, assimilating similar datasets individually has been shown to improve convective forecasts during the daytime (e.g., Kawabata et al. 2007; Wulfmeyer et al. 2006), primarily through modifications of the low-level moisture field. For example, Benjamin et al. (2004) and Kawabata et al. (2014) show that assimilating RWPs or Doppler lidars can enhance moisture advection and increase the convective available potential energy (CAPE) for developing convection. Additionally, Coniglio et al. (2019) recently perform the first evaluation of the impact of assimilating AERI data and find that the high-frequency thermodynamic retrievals can lead to small improvements in short-term convective forecasts.

Given that remote sensing instruments can provide high-frequency profiles of elevated features, we hypothesize that assimilating these data could be especially impactful for improving forecasts of nocturnal convection. Chipilski et al. (2020) investigate this hypothesis by assimilating remote sensing data for a bore-aided MCS during PECAN. They find that assimilating AERIs enhances moisture advection into the MCS, whereas kinematic profilers improve model biases in the structure of the simulated bore. In Chapter 3 of this dissertation, we perform a similar analysis to determine whether assimilating observations collected during PECAN can improve forecasts of a nocturnal CI event. We specifically evaluate the impact of assimilating these data on mechanisms important to the CI process, including the elevated moist layer and nocturnal LLJ.

Though the first section of this dissertation addresses the impacts of assimilating PECAN data for a single CI event, nocturnal convection can also be generated and supported by many mechanisms that occur at a variety of scales. Weckwerth et al. (2019) document many types of CI observed during PECAN including larger-scale frontal overrunning events, events forming near an MCS, and CI developing along a bore or density current. Each type of CI event is dynamically different and occurs within a different thermodynamic environment. As such, the benefits of assimilating remote sensing datasets are likely case- and environment-dependent. This wide spectrum of nocturnal convective events motivates Chapter 4 of this dissertation which systematically evaluates the impact of assimilating the PECAN dataset for a large set of nocturnal CI events.

While many studies note the positive impacts of assimilating remote sensing instruments, others find that these data can sometimes degrade forecast skill for certain events. For example, Chipilski et al. (2020) find that assimilating AERIs also introduces a positive bias in the simulated bore height despite improving the overall convective forecast. Coniglio et al. (2019) also discuss a few cases where assimilating AERI data reduces forecast skill for CI forecasts during the day. While forecast degradations during DA have been documented previously (e.g., Sun et al. 2012; Djalalova et al. 2016), such impacts are frequently found when assimilating remote sensing data (Bormann et al. 2016). These findings indicate that many unanswered questions remain about how to best assimilate the thermodynamic and kinematic profilers collected during PECAN.

Incorrectly assigned observation errors could be one potential source of forecast degradations when assimilating remote sensing data. In DA, the analysis state is a function of the background and observation error covariances (Kalnay 2003). If either set of errors are incorrect, the final analysis state will be suboptimal and subsequent forecasts can suffer. Much DA research focuses on improving background error covariances through ensemble- or hybrid-based methods (Whitaker and Hamill 2002; Buehner et al. 2010), while observation errors are typically assumed to be constant (Bormann et al. 2016). However, observation errors for remote sensing profilers can include large systematic or representation components that vary significantly in space and time (Fowler and Van Leeuwen 2013). As such, assuming a constant observation error when assimilating such data can naturally lead to some observations being underweighted (overestimated observation errors) or overweighted (underestimated observation errors) during DA.

Recent works evaluate the benefits of adaptively inflating observation errors for satellite data (e.g., Geer and Bauer 2011; Minamide and Zhang 2017). These studies find potential forecast improvements by limiting impacts of cloudy radiance data when the observation differs from the background forecast. Similar methods could potentially prove useful when assimilating AERI observations, as Turner and Blumberg (2019) document potential retrieval errors below cloud base due to issues separating cloud emission from the total measured emission. Desroziers et al. (2005) present another potential solution to estimating flow-dependent observation errors using observation-space diagnostics such as background and analysis residuals. Such methods have also been shown to be beneficial when assimilating satellite observations (e.g., Weston et al. 2014; Cambell et al. 2017), though these applications sacrifice time-dependent observation errors for statistical robustness. To further improve forecasts of nocturnal convection, Chapter 5 expands upon previous chapters by evaluating whether novel methods for assigning flow-dependent observation errors can further improve the benefits of assimilating the PECAN dataset.

1.2 Overview of this dissertation

This dissertation employs a Gridpoint Statistical Interpolation (GSI)-based Ensemble Kalman filter (EnKF) to evaluate the impact of assimilating PECAN observations for a variety of nocturnal convective events. We also employ the Weather Research and Forecasting (WRF) – Advanced Research WRF model (WRF-ARW) as the forecast component of this NWP system. The configuration of these systems is detailed in Chapter 2, along with an overview of the PECAN dataset. Given the experimental nature of many of the instruments employed during PECAN, we perform quality control and observation pre-processing steps that are also described in Chapter 2. This dissertation then addresses three primary scientific questions.

Chapter 3 first explores whether assimilating PECAN observations can improve simulations of a nocturnal CI event and its relevant initiation mechanisms. The CI of interest occurred near an elevated moist layer located just north of the intersection of the LLJ with a synoptic boundary. This setup is commonly observed during CI events in the Great Plains (Weckwerth et al. 2019). In addition to the convective impacts, we perform data denial experiments that assess the relative impact of each observation type (i.e., thermodynamic profilers, kinematic profilers, surface observations).

Next, Chapter 4 addresses the systematic impact of assimilating the PECAN observations across a larger set of nocturnal CI events. We select 13 cases observed during PECAN that capture the wide spectrum of mechanisms responsible for nocturnal CI, including frontal overrunning events, CI occurring near an MCS, bore- or outflow-generated events, and pristine CI events that develop without the influence of parent convective systems or obvious surface boundaries. To evaluate the forecasts over a large set of cases, we also develop an object-based algorithm for detecting and verifying forecasts of CI that is detailed within Chapter 4. This chapter also specifically assesses the impact of assimilating collocated datasets.

Finally, Chapter 5 of this dissertation explores whether forecasts of nocturnal convection can be further improved by applying novel methods for assigning observation errors. These methods, including both static and flow-dependent methods for computing the full observation error, are detailed within the methods section of Chapter 5. To evaluate these methods, we select a case from the systematic study presented in Chapter 4 in which assimilating PECAN observations degrades the convective forecast skill. A general summary and plans for future work are given in Chapter 6.

Chapter 2

Methods and system description

In each chapter of this dissertation, we assimilate both conventional and PECAN observations using a GSI-based EnKF recently extended for meso- and convective-scales. GSI performs the observation quality control and computes the observation-space statistics, while EnKF computes the background error covariance and performs the DA update. Additionally, we pair the GSI-based EnKF with the WRF-ARW model (Skamarock et al. 2008) as the forecast element of the DA system. Each component of this system is detailed below along with a summary of the assimilated datasets.

2.1 Description and configuration of the GSI-based ensemble Kalman filter (EnKF)

During DA cycling, the GSI-based EnKF uses observations ($\mathbf{y^o}$) to update a prior estimate of the model state ($\mathbf{x^b}$) into an analysis ($\mathbf{x^a}$). The EnKF improves upon other DA methods such as 3D-Var by sampling flow-dependent background error covariances ($\mathbf{P^b}$) from ensemble forecasts (Johnson et al. 2015; Houtekamer and Zhang 2016). For this dissertation, we utilize the ensemble square root filter (EnSRF; Whitaker and Hamill 2002) version of the EnKF wherein observations are serially assimilated to update the mean state following:

$$\mathbf{\overline{x}^a} = \mathbf{\overline{x}^b} + \mathbf{K}(\mathbf{y^o} - \mathbf{H\overline{x}^b}) \tag{2.1}$$

$$\mathbf{K} = \mathbf{P^b H^T}(\mathbf{H P^b H^T} + \mathbf{R})^{-1} \tag{2.2}$$

Here, **K** represents the Kalman gain, **R** represents the observation error covariance matrix, and **H** is the linearized observation operator performed by GSI. To update the ensemble

Table 2.1: List of covariance localization radii applied throughout this dissertation.

Observation type	Horizontal localization (km)	Vertical localization $[ln(P/P_{ref})]$
Conventional observations, PECAN rawinsondes, AERIs	700, 200	1.1, 0.55
Doppler lidars and RWPs	700, 200	0.20, 0.20
Radar observations	N/A, 200	N/A, 0.55

member perturbations (given by the prime symbols), EnSRF computes a modified Kalman gain ($\mathbf{\tilde{K}}$) following:

$$\mathbf{x'^a} = \mathbf{x'^b} - \mathbf{\tilde{K}Hx'^b} \tag{2.3}$$

$$\mathbf{\tilde{K}} = (1 + \sqrt{\mathbf{R}/(\mathbf{HP^bH^T} + \mathbf{R})})^{-1}\mathbf{K} \tag{2.4}$$

We refer to Whitaker and Hamill (2002) for additional details on the EnKF and EnSRF, including the methods used to compute the matrix products $\mathbf{P^bH^T}$ and $\mathbf{HP^bH^T}$.

Computational costs limit the number of ensemble members that can be used to sample $\mathbf{P^b}$. To treat these sampling errors, we use covariance localization following a Gaspari and Cohn (1999) function that limits observation impacts beyond a certain distance. The localization radii for conventional observations are set following Degelia et al. (2018). For the PECAN observations, we perform sensitivity tests to determine the localization radii that produce the highest fractions skill score (FSS; Roberts and Lean 2008) for the case study described in Chapter 3 (values shown in Table 2.1). Additionally, to treat errors associated with the misrepresentation of background errors, both a constant 15% (multiplicative; Whitaker and Hamill 2012), as well as a relaxation to prior spread of 95% (Whitaker and Hamill 2012) inflation methods are applied to the ensemble during every DA cycle. The constant inflation parameters smoothly decrease to $\sim$3% at the model top to avoid overdispersion (Johnson et al. 2015; Zhu et al. 2013). Finally, to maintain ensemble spread and quickly introduce observed storms into the analyses, random Gaussian noise with zero mean is added during each cycle and to every grid point for temperature (σ = 0.5 K), dewpoint temperature (σ = 0.5 K), and horizontal winds (σ = 0.5 ms^{-1}) following Wang and Wang (2017).

2.2 Configuration of the WRF-ARW forecast system

All simulations presented herein utilize version 3.7.1 of WRF-ARW. Following Johnson et al. (2017a), we first generate 40 ensemble members by downscaling members 1 through 20 of both the Global Ensemble Forecast System (GEFS; Wei et al. 2008) and Short-Range Ensemble Forecast (SREF; Du et al. 2014). These operational simulations initialize, and update the lateral boundary conditions, for our outer, CONUS domain with 12-km grid spacing (d01 in Fig. 2.1). We note that the native GEFS and SREF systems have horizontal resolutions of approximately 34 km and 16 km, respectively. After DA is complete on the outer domain, an inner, convection-permitting domain with 4-km grid spacing is nested within the mesoscale grid (d02 in Fig. 2.1). The simulations presented in chapter 4 also utilize a third domain that is described later. Each domain features 50 vertical levels on a stretched grid with a 50 hPa model top. The vertical grid spacing is approximately 200 m in the planetary boundary layer (PBL) and increases to 450 m at 500 hPa.

Fixed physical parameterization schemes are used for all 40 members during both the DA and forecast periods that follow Degelia et al. (2018; schemes listed in Table 2.2). These include the 1.5-order, local closure Mellor-Yamada-Nakanishi-Niino (MYNN; Nakanishi and Niino 2006) PBL scheme that Coniglio et al. (2013) find to be nearly unbiased for nighttime profiles of temperature and moisture. During DA, we utilize the WRF single-moment 6-class (WSM6; Hong and Lim 2006) microphysical parameterization scheme to be consistent with the observation operator for reflectivity in GSI (Johnson et al. 2015; Wang and Wang 2017). We use the Lin et al. (1983) scheme during the forecast period that Degelia et al. (2018) subjectively find performs well for CI forecasts. Finally, we also apply the Noah land surface model (Ek et al. 2003), Rapid Radiative Transfer Model for general circulation models (longwave radiation; Iacono et al. 2008), Goddard shortwave radiation (Tao et al. 2003), and the Grell and Freitas (2013) cumulus parameterization schemes. Given that the 4-km domain can explicitly resolve convective storms, we only utilize the cumulus parameterization on d01.

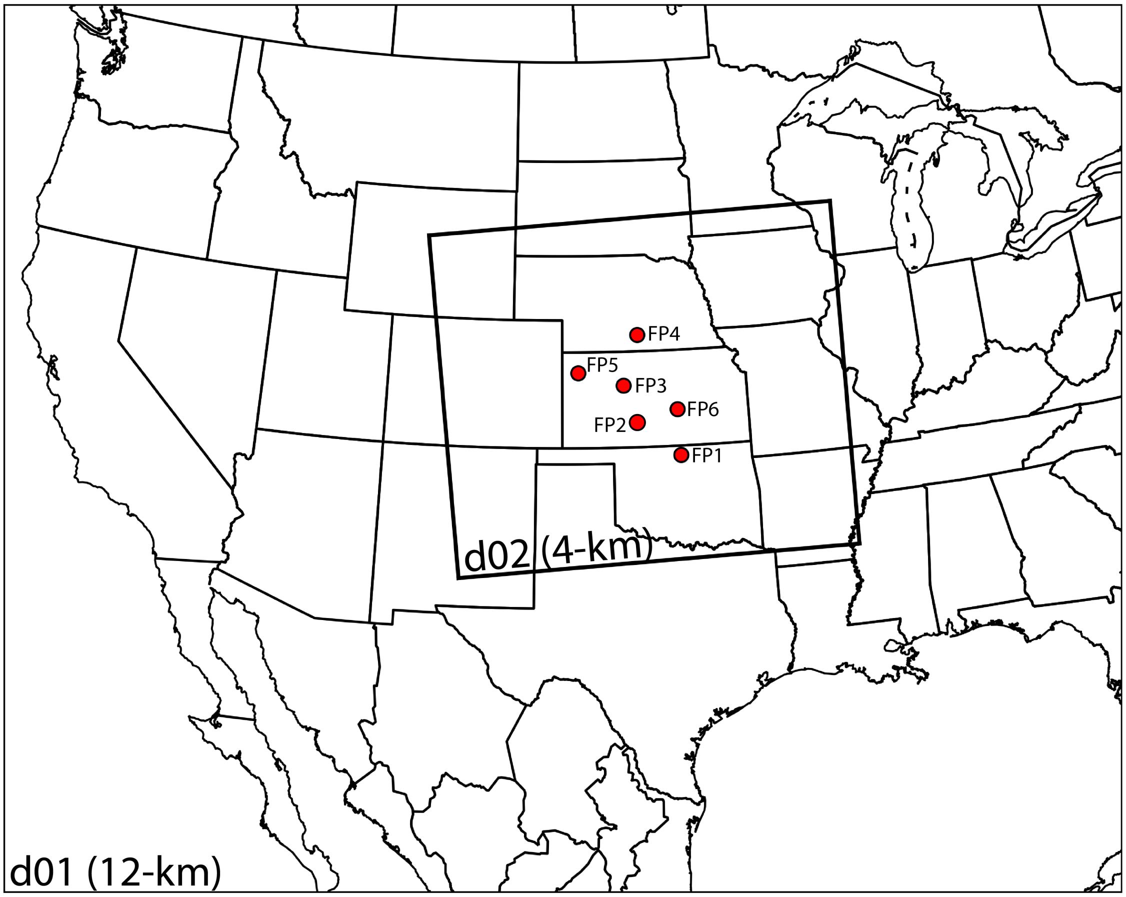


Figure 2.1: Domain configurations used during DA. Also shown are the location of each fixed PISA (FP) observing site.

Table 2.2: Physical parameterizations used in WRF-ARW.

Parameterization type	Scheme name
Microphysical parameterization (DA)	WSM6 (Hong and Lim 2006)
Microphysical parameterization (forecast)	Lin et al. (1983)
PBL parameterization	MYNN (Nakanishi and Niino 2006)
Longwave radiation parameterization	RRTMG (Iacono et al. 2008)
Shortwave radiation parameterization	Goddard (Tao et al. 2003)
LSM	Noah (Ek et al. 2003)
Cumulus parameterization	Grell and Freitas (2013)

2.3 Overview of assimilated datasets

Each section of this dissertation assimilates conventional, radar, and PECAN observations. The conventional data consist of operational surface, rawinsondes, aircraft, ship, and buoy observations obtained from the North American Mesoscale Data Assimilation System (NDAS; Rogers et al. 2009). The radar observations include reflectivity and radial velocity data that have been pre-processed using WDSS-II (Lakshmanan et al. 2007). We assimilate the full radar dataset (i.e., no thinning) from all WSR-88D stations within the PECAN domain, following the methods described in Johnson et al. (2015) and Degelia et al. (2018).

The PECAN observations (Table 2.3) are obtained from the PECAN field catalog (available online at http://catalog.eol.ucar.edu/pecan) and consist of data collected from both fixed and mobile PECAN Integrated Sounding Arrays (PISAs), mobile mesonets, and mobile GPS Advanced Upper-Air Sounding Systems (MGAUS). Each fixed (FP; Fig. 2.1) and mobile PISA (MP) features different instruments as described by Geerts et al. (2017). The instruments assimilated here are further described below but include AERIs (∼5-min thermodynamic profiles), Doppler lidars, RWPs, rawinsondes, and surface observations. The data are typically quality controlled before being uploaded to the field catalog, though we further pre-process each dataset using a variety of methods. The benefits these pre-processing steps are shown in Haghi et al. (2018). Additionally, we note that the exact instruments assimilated differ in each chapter.

2.3.1 AERIs

AERIs retrieve thermodynamic profiles from observations of downwelling infrared radiance (Turner and Löhnert 2014) and have been employed to study a variety of meteorological phenomena including the African monsoon (Hansell et al. 2010), shallow convective clouds (Fast et al. 2019), atmospheric bores (Toms et al. 2017), low-frequency gravity waves (Adams-Selin and Schumacher 2019), and nocturnal convection (Geerts et al. 2017). They can capture thermodynamic data on the order of minutes, allowing for the rapid monitoring

Table 2.3: List of PECAN observing sites and instruments that are assimilated across this dissertation. We note that the exact sites and instruments assimilated differ in each chapter. All AERI observations are produced by the AERIoe retrieval technique (Turner and Löhnert 2014) except at FP1 where the data are generated by AERIprof (Feltz et al. 2003). Additionally, we note that the FP1 site includes three RWPs spaced by an average of 17 km.

Site name	**Location**	**Instruments**	**Reference**
FP1	Lamont, OK	AERI (AERIprof) Doppler lidar 915 MHz wind profiler (3) Rawinsonde	Gero et al. 2014 Newsom and Krishnamurthy (2014) Sivaraman et al. (1990) UCAR/NCAR (2015a)
FP2	Greensburg, KS	AERI Rawinsonde Surface observations	Turner (2016a) Vermeesch (2015) Delgado and Vermeesch (2016)
FP3	Ellis, KS	AERI Doppler lidar 449 MHz wind profiler Rawinsonde Surface observations	Turner (2016b) Hanesiak and Turner (2016a) UCAR/NCAR (2017) Clark (2016) UCAR/NCAR (2015b)
FP4	Minden, NE	915 MHz wind profiler Rawinsonde Surface observations	UCAR/NCAR (2015c) UCAR/NCAR (2016a) UCAR/NCAR (2015b)
FP5	Brewster, KS	AERI 915 MHz wind profiler Rawinsonde Surface observations	Turner (2016c) UCAR/NCAR (2015d) UCAR/NCAR (2016b) UCAR/NCAR (2015b)
FP6	Hesston, KS	Doppler lidar Rawinsonde Surface observations	Hanesiak and Turner (2016b) Holdridge and Turner (2015) Turner (2016d)
MP1 "CLAMPS"	Mobile	AERI Doppler lidar Rawinsonde	Turner (2018) Turner (2016e) Klein et al. (2016)
MP2 "MIPS"	Mobile	915 MHz wind profiler	Knupp and Wade (2016)
MP3 "SPARC"	Mobile	AERI Doppler lidar Rawinsonde Surface observations	Wagner et al. (2016b) Wagner et al. (2016a) Wagner et al. (2016c) Wagner et al. (2016d)
MP4 "MISS"	Mobile	915 MHz wind profiler Rawinsonde Surface observations	UCAR/NCAR (2016c) UCAR/NCAR (2016d) UCAR/NCAR (2016c)
M-GAUS1	Mobile	Rawinsonde Surface observations	Ziegler et al. (2016) Waugh and Ziegler (2017)
M-GAUS2	Mobile	Rawinsonde Surface observations	Ziegler et al. (2016) Waugh and Ziegler (2017)
M-GAUS3	Mobile	Rawinsonde Surface observations	Ziegler et al. (2016) Waugh and Ziegler (2017)
MM1	Mobile	Surface observations	Waugh and Ziegler (2017)
MM2	Mobile	Surface observations	Waugh and Ziegler (2017)
NSSL-NOXP	Mobile	Surface observations	Waugh and Ziegler (2017)

of elevated mixed layers (Feltz et al. 2003) and boundary layer stability (Blumberg et al. 2017). Recently, Turner and Löhnert (2014) introduce an optimal estimation retrieval algorithm (AERIoe) that allows for computing unique uncertainty profiles for AERI observations, improving the potential impact of assimilating this dataset by allowing for less-confident retrievals, such as those in cloudy conditions, to be weighted less.

In this dissertation, we assimilate the temperature and moisture profiles output by the AERI retrieval (as opposed to the raw radiance observations). Because the retrieval accuracy quickly decreases with height and above cloudy layers, no observations above either 3 km AGL or cloud base are assimilated here (D. Turner, personal communication). Additionally, to reduce both the correlated and uncorrelated observation errors, a "superob" method (e.g., Berger 2004) is applied to each retrieval wherein the observations are averaged over a 10 hPa depth. For all observation types, we only apply the superob method to the portions of the profile where the native observation spacing is less than the superob depth. We also do not apply any temporal averaging or thinning to the AERI data, as one large advantage of assimilating this instrument is its ability to capture rapid changes in moisture (Chipilski et al. 2020).

2.3.2 Vertical wind profilers

Multiple types of vertical wind profilers, including both RWPs and Doppler lidars, were deployed during PECAN. These instruments are also commonly used for both forecast and research applications. Similar to Doppler weather radars, RWPs actively observe horizontal winds by measuring the radio waves produced by Bragg scattering from inhomogeneities in the thermodynamic field. Doppler lidars operate similarly, measuring light waves backscattered by smaller aerosols. These kinematic profilers can be used to measure various phenomena that conventional observations typically cannot observe with high vertical or temporal resolution, including LLJs (Banta et al. 2002; Smith et al. 2019) or atmospheric turbulence (Smalikho et al. 2005; Calhoun et al. 2006; O'Connor et al. 2010).

Except at FP3, the Doppler lidars were not co-located with RWPs. Thus, RWPs and Doppler lidars could often be considered complimentary. However, differences in the design of the instruments could lead to different DA impacts. Doppler lidars are only able to collect useful information from the lowest 1-3 km of the atmosphere due to the depth of potential scatterers (Menzies and Hardesty 1989). Doppler lidars also tend to sample finer-scale flow fields, such as turbulence, compared to RWPs. Conversely, RWPs collect backscatter from larger particles (e.g., hydrometeors, dust, insects) which are present as high as 10 km above ground level (AGL).

Here, we only assimilate vertical profiles of zonal and meridional wind (i.e., no radial velocity data) collected by these wind profilers. The vertical profiles from each RWP are provided as 30-min averages. At FP4, FP5, and MP4, the RWPs are calculated using the Improved Moments Algorithm (Morse et al. 2002) which provides a confidence measure for each observation. We reject any data with a confidence below 0.5 as recommended by the data providers. At FP3, any 449 MHz wind profiler data with a signal-to-noise ratio less than -6 *dB* are rejected (W. Brown, personal communication). Furthermore, at FP4 and FP5, the 915 MHz profilers operate in both a "low" mode which features 180 m vertical sampling up to 4 km AGL, and a "high" mode which features 360 m vertical sampling up to 12 km AGL. We choose to form a composite profile at these sites by rejecting any "high" mode data below 4 km AGL. The superob method with a depth of 100 m (similar to a depth of 10 hPa in the boundary layer) is again applied to these observations, as no pressure data are provided.

To remove the impacts of turbulence not resolved by our model and to be consistent with the averaging window for RWPs, we also temporally average Doppler lidar observations into 30-min profiles. Data below 100 m or above 3000 m AGL are not assimilated due to quality issues (D. Turner and P. Klein, personal communication). We also perform gross checks to remove any erroneous data based on the root-mean-square difference between the observed radial velocity and its fitted values. Again, the lidar wind observations are superobbed to a depth of 100 m.

2.3.3 Rawinsondes and mobile surface observations

During PECAN, rawinsondes and surface data were collected at much higher frequency than operational networks. The PECAN rawinsondes were launched as often as every 3 h and at non-standard times. Because Privé et al. (2017) show significant improvements when assimilating rawinsondes more frequently, we expect the assimilation of PECAN rawinsondes to produce a larger impact than those demonstrated by previous studies that only evaluate the operational network (e.g., Benjamin et al. 2010). The rawinsonde data are provided with quality control markers following the methods described in Loehrer et al. (1996). We only assimilate levels where all data are marked as ”good”. The rawinsonde data are also superobbed to a depth of 10 hPa to be consistent with the other PECAN observations.

Lastly, the PECAN surface observations were typically collected every 1-5 min at most PECAN platforms. These data, consisting of temperature, humidity, wind, and pressure, are thinned to 5-min intervals due to the high frequency of the data and to limit potential model shock. Other than gross checks for valid data, no other quality control methods are applied to the surface observations before assimilation.

Chapter 3

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 PECAN

3.1 Introduction

CI represents the process through which air parcels are successfully lifted beyond their level of free convection (LFC) and buoyantly accelerate upwards to produce a deep and precipitating updraft (Crook and Klemp 2000; Kain et al. 2013). Predicting the timing and location of CI remains one of the most difficult tasks in meteorology (Lilly 1990), as many studies demonstrate how small initial condition perturbations within the range of observational uncertainty can significantly modify forecasts (Crook 1990; Martin and Xue 2006). At night in the Great Plains of the United States, CI commonly contributes to the nocturnal maximum in summer precipitation (e.g., Surcel et al. 2010; Weckwerth et al. 2019) and leads to thunderstorms which produce all severe weather hazards (Reif and Bluestein 2017). Despite these impacts, predicting CI at night can be even more of a challenge compared to its daytime counterpart (Keclik et al. 2017).

Past studies show that NWP models that employ convective parameterizations underpredict nocturnal CI events in the Great Plains (Davis et al. 2006). Although various deficiencies have been resolved through the use of convection-allowing models (Weisman et al. 2008), modern NWP still struggles to predict many of the mechanisms that initiate convection at night. As discussed in Chapter 1, NWP models routinely fail to correctly forecast the LLJ (Storm et al. 2009; Shin and Hong 2011; Smith et al. 2015; Johnson and Wang 2017; Johnson et al. 2017a) and atmospheric bores (Johnson et al. 2017b; Chipilski et al. 2020). Additionally, elevated moisture biases at night are common and can lead to significant errors in CI forecasts. Peters et al. (2017) correlate errors in elevated moist layers with simulations of highly impactful MCSs. They show that simulations with negative moisture biases produce large errors in both the timing and location of MCS initiation due to parcels requiring additional residence time within the lifting regions.

Recently, Degelia et al. (2018) show improvements to a nocturnal CI forecast by assimilating conventional and radar observations. They find that assimilating these data enhances the buoyancy and convergence prior to CI, while the radar observations aid in suppressing spurious convection and erroneous outflow boundaries. However, the observations assimilated in Degelia et al. (2018) have become routinely assimilated in operational centers and their impacts are now relatively understood. This chapter expands upon the findings of Degelia et al. (2018) by evaluating the forecast impact of assimilating the novel dataset collected during PECAN (Geerts et al. 2017). The PECAN project seeks to better understand the processes responsible for nocturnal convection in the Great Plains with a focus on nocturnal CI, MCSs, atmospheric bores, and the LLJ (Geerts et al. 2017). The data collected during the field campaign included a network of thermodynamic and kinematic profilers similar to those recommended by the National Research Council (2009).

The observations assimilated here consist of AERIs, Doppler wind lidars (e.g., Menzies and Hardesty 1989), RWPs (e.g., Benjamin et al. 2004), high-frequency rawinsondes, and special surface data collected from fixed and mobile PECAN platforms. Assimilating similar datasets has been shown to improve convective-scale forecasts of various features (e.g., Kawabata et al. 2007; Wulfmeyer et al. 2006). Benjamin et al. (2004) and Kawabata et al. (2007) demonstrate forecast improvements when assimilating kinematic profilers, and Hitchcock et al. (2016) show midlevel moisture impacts when assimilating special rawinsonde observations. Until recently, little work has evaluated the impact of assimilating ground-based thermodynamic profilers. A study by Coniglio et al. (2019) shows that assimilating high-frequency thermodynamic retrievals collected by AERIs can lead to improvements, albeit non-significant, in short-term convective forecasts. However, the

Coniglio et al. (2019) study only assimilates data from a single AERI, and for a short period (2-5 hours) prior to CI. Therefore, we aim to expand upon previous works by assimilating data collected by multiple AERI platforms and over a longer period of assimilation.

This dissertation chapter focuses on the 26 June 2015 nocturnal CI event during PECAN. The CI of interest occurred near an elevated moist layer located just north of the intersection of the LLJ with a synoptic boundary. Such placement is commonly observed during nocturnal CI events in the Great Plains (Weckwerth et al. 2019). This chapter tests the hypothesis that assimilating a large network of many different PECAN observations can improve the simulation of both the elevated moist layer and the ascent mechanisms. In addition to evaluating the impact of assimilating the entire PECAN dataset, data denial experiments are presented to assess the relative impact of each observation type.

An overview of the 26 June 2015 CI is presented in section 3.2. Section 3.3 discusses the cycling configuration and observation errors used to assimilate PECAN observations for the CI case study. The assimilation and data denial results are found in section 3.4. An ingredients-based approach is applied in section 3.5 to better understand what aspects of the environment lead to the observation sensitivities for CI, while we explore the moisture and kinematic impacts during the DA cycling in section 3.6. A final discussion is found in section 3.7.

3.2 Overview of the 26 June 2015 nocturnal CI event

As an upper-level ridge deepened over the southwestern United States on 25 June 2015 (Fig. 3.1a), northwesterly flow developed above the Central Plains. A surface low, related to an embedded shortwave trough (Fig. 3.1a), strengthened a stalled, pre-existing frontal boundary into a synoptic cold front (Fig. 3.1a,b). Additionally, a cold pool generated by early-afternoon convection appears to have further reinforced this synoptic front (Fig. 3.2b). By the late afternoon of 25 June, stronger, surface-based cells developed along the synoptic boundary in central Kansas (Fig. 3.2c). After sunset at 0154 UTC (2054

LST), a southwesterly, criterion-1 LLJ (12.5 ms^{-1}; Bonner 1968) developed across western Oklahoma and central Kansas (Fig. 3.1b).

Conditions were favorable for further convective development after sunset on 25-26 June. First, large-scale isentropic ascent developed throughout northern Kansas due to the interaction of the LLJ with the synoptic boundary (Fig. 3.1b). Second, an additional mesoscale convergence zone associated with the northern terminus of the LLJ was present in northeastern Kansas (circled in Fig. 3.1b). At approximately 0215 UTC 26 June, a linear band of convective cells, which were disconnected from storms along the synoptic boundary, initiated in northern Kansas. These cells began to merge with additional clusters of convection that developed in northwestern Missouri (Fig. 3.2g). This arcing band of nocturnal convection (Fig. 3.2e,f) is the focal point of this chapter. The convective cluster continued to grow upscale into an MCS (Fig. 3.2h) that propagated southeastward, producing both severe wind and flash flooding throughout eastern Kansas. We note that many other nocturnal CI events occurred throughout Kansas on 26 June 2015, some of which are discussed in Trier et al. (2017).

Mobile observing platforms were deployed for this event as part of intensive observing period (IOP) 16. A sounding taken by a PECAN vehicle showed a moist layer atop the frontal inversion north of the synoptic boundary (Fig. 3.1c). Significant instability (>3000 $J\ kg^{-1}$ of CAPE) was associated with elevated air parcels, although some inhibition had to be overcome before CI could take place. As surface-based parcels were located below the frontal inversion, much of the nocturnal CI episode of interest was likely elevated. However, recent analyses of this event by Trier et al. (2018) and Sun and Trier (2018) highlight the potential role of outflow boundaries in the southern portion of the nocturnal CI event. In particular, the surface-based cells in Fig. 3.2c produced an outflow boundary that moved northward through the region of nocturnal CI and that will be discussed throughout this text. These findings indicate that some of the early cells in this CI event (Fig. 3.2e) were surface-based.

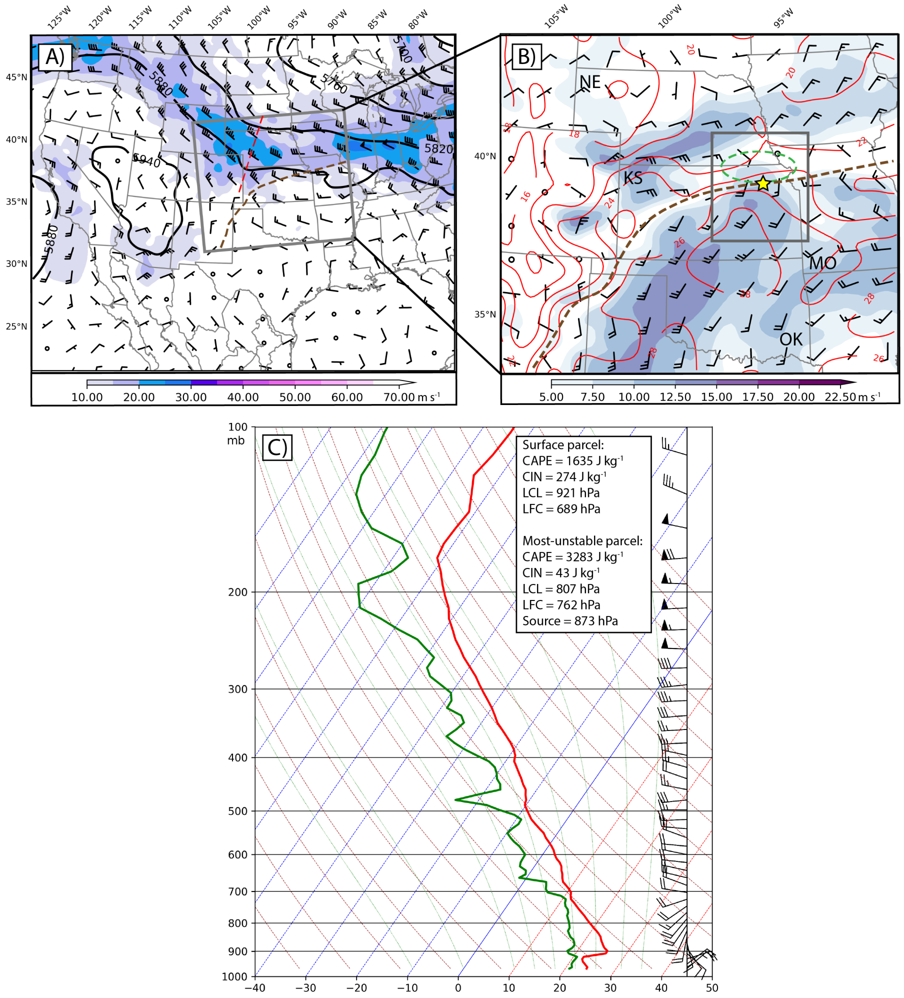


Figure 3.1: Overview of the mechanisms leading to nocturnal CI on 26 June including (a) 500 hPa geopotential height (m; black contours) and winds (ms^{-1}; fill and barbs) valid at 0300 UTC 26 June; (b) 850 hPa winds (ms^{-1}; fill and barbs) and surface temperature contours (°C; red); and (c) a PECAN sounding taken from the location of the yellow star in (b) at 0215 UTC 26 June. In (a,b) the half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. The red dashed line in (a) represents the location of a shortwave trough axis, the brown dashed line in (a,b) represents the location of the synoptic boundary discussed in the text, and the dashed green circle in (b) indicates the approximate LLJ terminus. The plotting domains in (a,b) represent the outer and inner domains used for the simulations in this chapter, respectively. The gray box in (b) indicates the plotting domain for 3.8. The states of Kansas (KS), Nebraska (NE), Missouri (MO), and Oklahoma (OK) are also labeled in (b). The plots in (a,b) are created using model analyses from the Rapid Refresh model (Benjamin et al. 2016). The sounding in (c) is plotted using the quality-control checks provided by Loehrer et al. (1996). Except for the wind data being superobed to a depth of 20 hPa, the data in (c) are processed using the same methods described in Chapter 2.

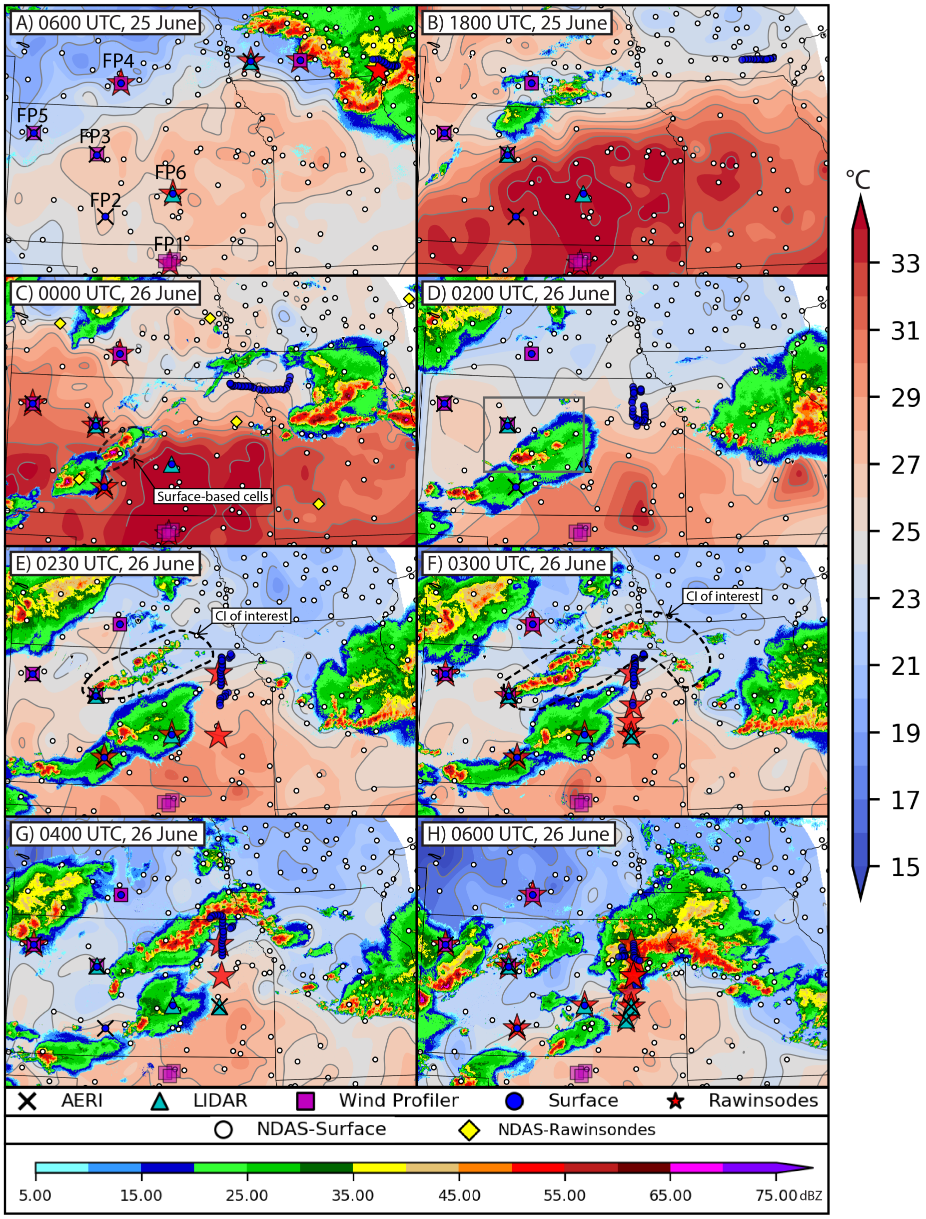


Figure 3.2: Evolution of the 25-26 June nocturnal CI event in terms of composite reflectivity (dBZ; bottom color bar). The reflectivity data are provided by the Multi-Radar Multi-Sensor dataset (MRMS; Smith et al. 2016). Additionally, an objective analysis of surface temperature (°C; right color bar) is shown using a two-dimensional linear spline interpolation. Surface temperature contours are also plotted in gray every 2 °C. The observations in the objective analysis are obtained from the MesoWest program (Horel et al. 2002). The fixed (FP) and mobile (MP) PECAN Integrated Sounding Arrays (PISAs) are labeled in (a). Also overlaid are both the PECAN observations and conventional NDAS observations that are assimilated in this chapter. The circle in (c) indicates the location of two surface-based cells discussed in the text. The gray box shown in (d) indicates the plotting domain for Fig. 3.11. The circles in (e,f) indicate the nocturnal CI event of interest.

3.3 Methods

3.3.1 Observation error treatment

In any DA algorithm, the observation error covariance matrix partially controls the weight between the observation and background states. When assimilating a new dataset, effort should be paid to how the observation errors are diagnosed (Bormann et al. 2011). Because RWPs, rawinsondes, and surface observations are routinely assimilated in operational systems, we assimilate those PECAN observations using a pre-existing, static error profile built into the GSI software (Shao et al. 2016). Conversely, both AERIs and Doppler lidars are considered experimental and thus their observation errors are less understood. Luckily, both the AERIoe and lidar algorithms provide unique error profiles for each observing time using the methods in Turner and Löhnert (2014) and Newsom and Krishnamurthy (2014), respectively. Assimilating these novel observations with unique error profiles allows for less confident retrievals to have a lower weight in the analysis.

In addition to the instrument error, the observation errors used in a DA system should also include contributions from representation errors (Geer and Bauer 2011). To account for these and any other residual errors, we inflate the AERI and lidar observation error profiles using:

$$\sigma_{Pf}^2 = \sigma_{Pi}^2 + \alpha(\sigma_{Sf}^2 - \sigma_{Si}^2) \tag{3.1}$$

where σ_{Pi}^2 is the instrument observation error variance profile provided by the PECAN dataset and σ_{Pf}^2 is the final observation error variance profile used for DA. The $(\sigma_{Sf}^2 - \sigma_{Si}^2)$ term represents an initial estimate of the residual error profile for profiling instruments based on the difference between an instrument uncertainty profile for rawinsondes (σ_{Si}^2; provided by Vaisala 2017) and the full error profiles for assimilating rawinsondes in GSI (σ_{Sf}^2). The static error profiles in terms of σ_{si} and σ_{sf} are shown in red in Fig. 3.3. The initial estimate of residual error is then tuned using the parameter α that varies by instrument. The values of α are chosen by comparing the skill of nocturnal CI forecasts when varying α by intervals of 0.25. The selected values are annotated in Fig. 3.3. Through a trial-and-error process, we find improved forecasts when linearly increasing α with height for AERI observations, such that observation errors near the top of the profile are inflated more. We hypothesize that this is because the observations errors output by AERIoe only include the diagonal terms of its posterior error covariance matrix, whereas the off-diagonal terms are shown to increase with height in Turner and Löhnert (2014). Additionally, as will be discussed in Chapter 5, the need for increased inflation at the top of the retrieval could be a result of increased uncertainty below cloud layers.

Example profiles of input and inflated observation errors for AERIs and Doppler lidars are shown in Fig. 3.3. Using this method, the forecast skill for the 26 June nocturnal CI event is improved compared to assimilating these observations using rawinsonde errors (not shown). Geer and Bauer (2011) and Minamide and Zhang (2017) use a similar approach to inflate observation error covariances for microwave imager radiances. We note that this technique is only meant as a preliminary method for assimilating the AERI and Doppler lidar observations. In Chapter 5, we develop and test additional methods for determining observation errors for these instruments.

3.3.2 Experimental design and configuration of DA cycling

To evaluate the impact of assimilating PECAN observations on the nocturnal CI forecast, we compare an experiment with all IOP observations assimilated (ALL) against a baseline forecast that only assimilates radar and conventional data (DENYALLPECAN). Additionally, we evaluate the relative forecast impact of AERIs, Doppler lidars, RWPs, rawinsondes, and mobile surface observations (see observation locations in Fig. 3.2) through data denial experiments (Table 3.1). In the data denial framework, a decrease in forecast skill in a denial experiment indicates a positive impact when assimilating those specific observations. For the experiments here, the observations are denied from the assimilation on both the outer and inner domain.

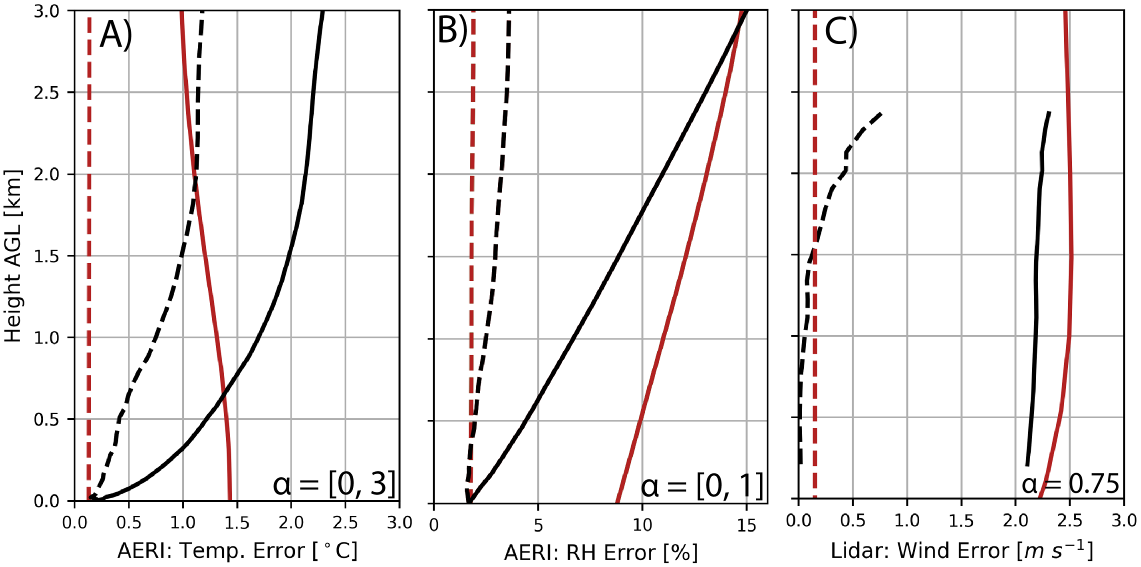


Figure 3.3: Observation error profiles for (a) AERI temperatures (°C); (b) AERI moisture (%RH); and (c) Doppler lidar winds (ms^{-1}) used in Eq. 3.1. The plots are calculated by averaging each individual error profile associated with the respective AERI or Doppler lidar observations assimilated in this study. The red lines represent the rawinsonde instrument error obtained from Vaisala (2017; σ_{si}; dotted) and the static rawinsonde error from GSI (σ_{sf}; solid). The black lines represent the input observation error profile provided by the PECAN dataset (σ_{Pi}; dotted) and the final observation error profile used for DA after inflation (σ_{Pf}; solid). The values used for α in Eq. 3.1 are also annotated for each variable. For AERI observation errors, the values of α vary linearly with height (from zero to the final bracketed value at 3 km AGL).

Table 3.1: List of experiments in Chapter 3.

Experiment name	Observations assimilated
DENYALL	Conventional and radar observations
ALL	Conventional, radar observations, and all PECAN observations listed in Table 2.3
DENYAERI	All observations from ALL except for AERIs
DENYLIDAR	All observations from ALL except for Doppler lidars
DENYWPROF	All observations from ALL except for RWPs
DENYSONDE	All observations from ALL except for PECAN rawinsondes
DENYSFC	All observations from ALL except for PECAN surface observations

The cycling description that follows describes the ALL experiment. On the outer domain, conventional data (see Fig. 3.2) are assimilated at 3-h intervals from 0000 to 2100 UTC 25 June. While the assimilation interval is 3-h, only observations from a 1-h time window (±30 min centered on the analysis time) are assimilated on the outer domain. We choose to also assimilate PECAN observations on the outer, mesoscale domain with the same cycling configuration as the conventional observations. This is because elevated moist layers, which occur on the mesoscale and are often associated with nocturnal CI (Wilson et al. 2018), could likely be improved by assimilating the thermodynamic profilers located at the FP sites. After 2100 UTC, the inner domain is initialized within the outer domain and conventional, PECAN, and radar observations are assimilated at 10-minute cycling intervals from 2110 UTC 25 June to 0000 UTC 26 June. As the WSR-88D network sufficiently covers the domain of interest, we choose not to assimilate any special radar data collected by PECAN instruments. After the final DA cycle at 0000 UTC, 7.5-h forecasts are initialized from members 1 through 20 of the DA ensemble to cover the nocturnal CI event.

3.4 Overview of the forecast results when assimilating the PECAN dataset

Before discussing the forecast results from individual experiments, we present the consistency ratio (Dowell et al. 2004) for each experiment in Fig. 3.4d-f. The consistency ratio acts as an evaluation of the analysis system by calculating the ratio between the square of the total ensemble spread in observation space (Wheatley et al. 2014) and the root-mean-square innovation. A value of 1.0 indicates that the ensemble spread fully accounts for the background ensemble error, while values greater or less than 1.0 indicate overdispersion or underdispersion, respectively. During the outer domain DA cycles, the consistency ratio for thermodynamic variables remains less than one for each experiment (Fig. 3.4d-f), indicating that the ensemble spread is not sufficient to represent the background ensemble errors. After downscaling to the inner domain (Fig. 3.4d-f), the consistency ratios for the thermodynamic

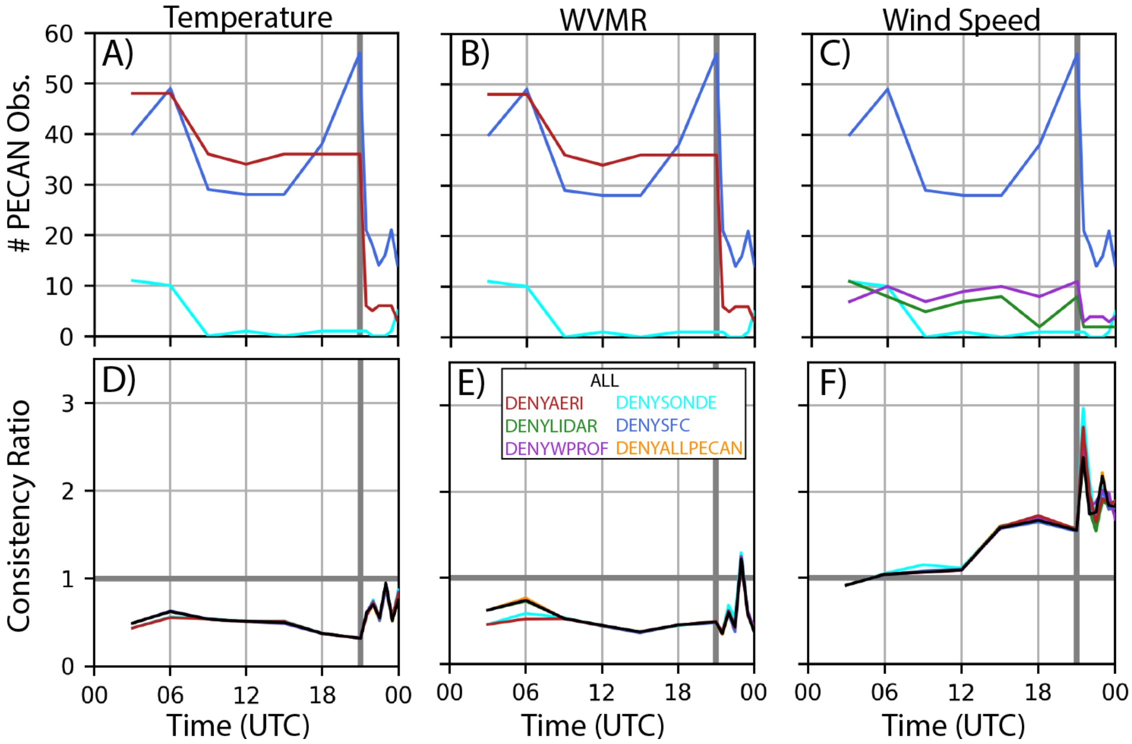


Figure 3.4: Series of (a-c) the number of PECAN observations assimilated in the ALL experiment; and (d-f) consistency ratio at each data assimilation cycle. The vertical gray lines indicate the transition from assimilating the PECAN observations on the outer domain (every 3 hours) to assimilating them on the inner domain (every 10 minutes). The inner domain observation counts and consistency ratios (right of the vertical gray lines) are only plotted for every third cycle. The colors in (a-c) correspond to the observation type assimilated in ALL that was denied from its respective denial experiments labeled in (d-f). For example, the red line in (a-c) represents the number of AERI observations assimilated in ALL.

variables are closer to 1.0. For the wind speed, the ensembles are typically overdispersive during both the outer and inner domain assimilation periods (Fig. 3.4f). Nevertheless, the ensemble statistics show no sign of filter divergence and each experiment produces generally similar values. Therefore, we assume that the DA system performs well enough for comparisons between the denial experiments.

3.4.1 Comparisons with an operational forecast of nocturnal CI

To first demonstrate the forecast impact of assimilating the PECAN dataset, we compare ALL and DENYALLPECAN with an operational forecast from the High-Resolution Rapid Refresh (HRRR; Earth System Research Laboratory 2016) model initialized at 0000 UTC 26 June. The forecast results from the HRRR are representative of other real-time simulations the 26 June nocturnal CI event. The experiments are first compared through raw neighborhood ensemble probability (NEP; Fig. 3.5) calculated using an 8-km neighborhood (Schwartz and Sobash 2017). Because the HRRR is a deterministic forecast as opposed to ensemble-based, Fig. 3.5a-d are presented as neighborhood probabilities (NP), which are equivalent to the NEP calculated with a single ensemble member. Additionally, the forecasts are compared through a time-series of FSS (Schwartz et al. 2010) calculated using 1-h accumulated precipitation data with a threshold of 2.54 $mm\ hr^{-1}$ (Fig. 3.6). The gridded precipitation data are provided by the Multi-Radar Multi-Sensor project (MRMS; Zhang et al. 2016) at 1-km resolution before being interpolated onto our 4-km forecast domain. The FSS is calculated over the box shown in Fig. 3.5l to ensure verification only over the event of interest. By using NEP as the input for calculating FSS (Schwartz et al. 2010), the score represents an ensemble verification metric (FSS for the HRRR is calculated using NP instead).

Even though the nocturnal CI of interest was likely at least partially driven by the large-scale mechanisms discussed in section 3.2, the real-time HRRR simulations largely fail to capture the event (Fig. 3.5a-d). Between 0300 and 0400 UTC, the HRRR generates convection too far north and does not produce a southwest-northeast oriented linear event as was observed. The FSS for the HRRR rapidly decreases from ∼0.45 to 0.05 (Fig. 3.6) during this time period. While the simulations capture fairly well the convection forming in western Missouri (Fig. 3.5b,c), the HRRR only generates weak probabilities in southeastern Nebraska that fail to match the observed locations or orientation of the nocturnal CI event.

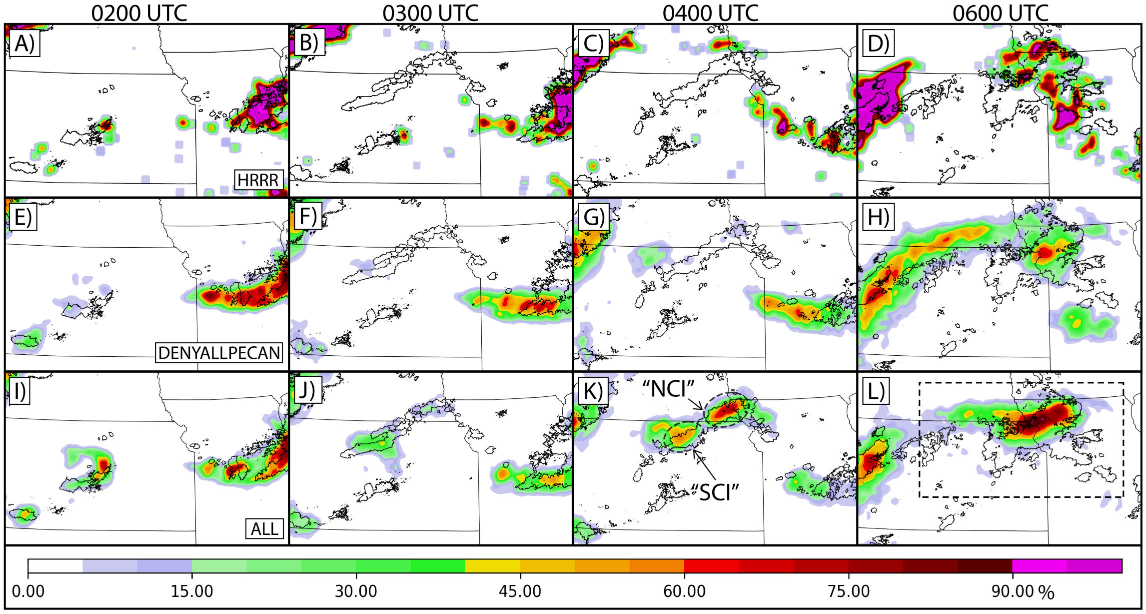


Figure 3.5: Probability forecasts of composite reflectivity greater than 30 dBZ for (a-d) the operational HRRR forecast initialized at 0000 UTC 26 June; (e-h) the DENYALL experiment; and (i-l) the ALL experiment. The contours of observed composite reflectivity greater than 30 dBZ are also overlaid in black. The probabilities for (a-d) are calculated as the neighborhood probabilities with a radius of 8 km, while the probabilities in (e-l) are calculated as neighborhood ensemble probabilities with a radius of 8 km. The northern (NCI) and southern (SCI) initiation episodes are annotated in (k). The black box in (l) indicates the domain used to calculate fractions skill score shown in Fig. 3.6.

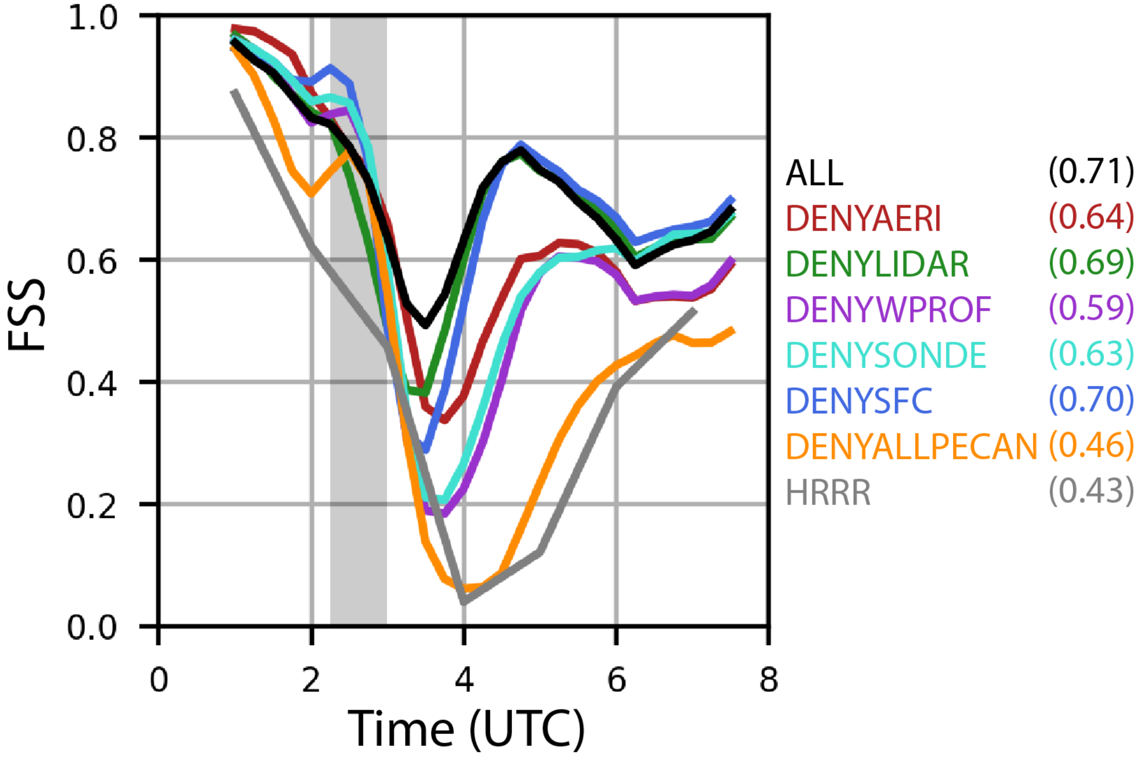


Figure 3.6: Fractions skill score (FSS) calculated with an 8-km neighborhood over the black box shown in Fig. 3.5l for a 2.54 $mm\ hr^{-1}$ precipitation threshold. The gray shading represents the approximate time of CI (between 0215-0300 UTC). The values shown in parenthesis signify the FSS when averaged over the entire forecast period.

DENYALLPECAN demonstrates similar issues to the HRRR (Fig. 3.5e-h) and the FSS for the two forecasts are similar (Fig. 3.6). Again, the linear event is almost entirely missed apart from low probabilities of two convective events at the extreme ends of the line at 0400 UTC (Fig. 3.5g). These signals are not maintained and do not merge into a linear cluster. Eventually, DENYALLPECAN generates a new linear system, but it forms further west than the observed event and is likely associated with a second CI event that is discussed in Trier et al. (2017). As the DENYALLPECAN experiment performs poorly and similar to the HRRR simulations, we assume that it serves as an accurate baseline to measure the advances that could be made when assimilating PECAN observations in an operational setting. We note that the lifting mechanisms discussed later are captured in both the HRRR and DENYALLPECAN (not shown). Thus, we hypothesize that the issues with these forecasts are primarily related to biases in the elevated instability profile.

Large forecast improvements are made when assimilating the IOP observations in ALL (Fig. 3.5i-l). The FSS for ALL first becomes larger than DENYALLPECAN at 0115 UTC (Fig. 3.6) due to improvements in resolving the ongoing surface-based convection in central Kansas. Shortly before 0300 UTC, ALL generates two distinct CI episodes along the northern and southern edge of observed linear event, henceforth called NCI (northern CI) and SCI (southern CI), respectively (Fig. 3.5k). By 0400 UTC, NCI and SCI congeal into a single linear event that closely matches the position and extent of the observed 30 dBZ contours (Fig. 3.5k). The linear convection in ALL then merges with additional convection in western Missouri to grow into a larger MCS, as was observed. Although the shape of the later MCS is not precisely captured in ALL, the experiment correctly predicts a strongly organized MCS along the northern Kansas and Missouri border by 0600 UTC (Fig. 3.5l). Figure 3.6 demonstrates that after CI is simulated at 0300 UTC, ALL maintains higher skill than DENYALLPECAN throughout the entirety of the forecast.

3.4.2 Data denial experiments

Data denial experiments based on ALL are used to determine the relative impact of each individual PECAN observation type on the forecasts of SCI and NCI. The same NEP plots from Fig. 3.5 are presented for the denial experiments in Fig. 3.7 and the skill scores for each are shown in Fig. 3.6. Before discussing the experiments individually, note that ALL simulates NCI and SCI, as well as the upscale growth into an MCS, better than any denial experiment. The FSS is also higher in ALL than the other experiments shortly after CI (Fig. 3.6), indicating that all the individual observation types in the PECAN dataset have a positive impact on the CI forecast.

Prior to CI, DENYAERI performs slightly better than ALL (Fig. 3.6) due to it better capturing the decaying surface-based convection in central Kansas (Fig. 3.7e). However, large improvements from the assimilation of AERI observations appear shortly after 0300 UTC when ALL correctly begins to generate NCI in Nebraska (Fig. 3.7b). DENYAERI does not produce the same convective cluster until 0330 UTC (not shown) and the FSS values are reduced from 0.50 in ALL to 0.35 in DENYAERI (Fig. 3.6). Because the convection within NCI eventually grows upscale into an MCS, DENYAERI also forecasts lower NEP values and produces a smaller MCS than ALL (Fig. 3.7d,h). These impacts are demonstrated in Fig. 3.6 as well, as the FSS for DENYAERI becomes lower than ALL after CI occurs and remains that way throughout the forecast period (Fig. 3.6). Thus, assimilating the AERI observations leads to a positive forecast impact for the northern cluster of CI and the later MCS.

Assimilating Doppler lidar observations has a smaller impact compared to the other observation types evaluated in this chapter. DENYLIDAR forecasts NCI similar to ALL. The only apparent differences for NCI are that DENYLIDAR performs slightly better at capturing the additional convective events forming along the LLJ terminus in far western Missouri (Fig. 3.7b,j). Instead, DENYLIDAR shows a small decrease in skill around 0300 UTC (Fig. 3.6) that is primarily connected to the reduced probabilities and extent of SCI. At

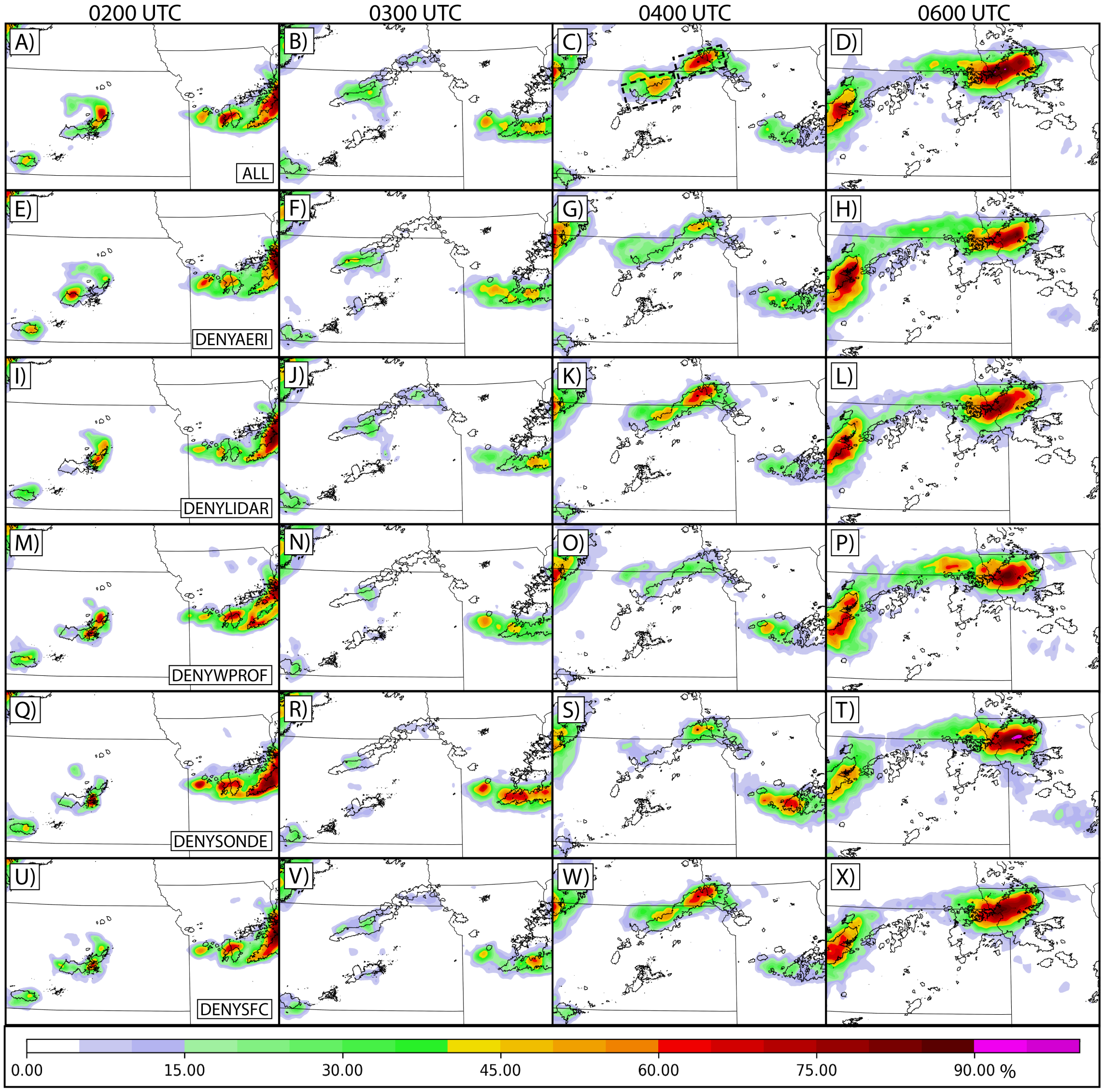


Figure 3.7: Neighborhood ensemble probability forecasts for (a-d) ALL; (e-h) DENYAERI; (i-l) DENYLIDAR; (m-p) DENYWPROF; (q-t) DENYSONDE; and (u-x) DENYSFC. Each plot is calculated with a radius of 8 km. The dashed black boxes in (c) indicate the averaging regions for the profiles in Figs. 3.10-3.11.

0300 UTC, the maximum NEP values for SCI are reduced by $\sim$15% in DENYLIDAR (Fig. 3.7b,j) compared to ALL. However, these differences in FSS do not remain large after the convection grows upscale (Fig. 3.7d,l).

Similar to the AERI observations, assimilating RWPs in ALL improves the forecast timing of NCI (Fig. 3.7b,n). Like DENYAERI, the convection that forms in southeastern Nebraska is not simulated by DENYWPROF until 0330 UTC (not shown). However, DENYWPROF also poorly captures SCI. Without assimilating the wind profiler data, the NEP values for SCI are reduced by nearly 40% compared to ALL at 0400 UTC (Fig. 3.7c,o). These large benefits are maintained throughout the upscale growth of the convective episodes into an MCS (Fig. 3.7d,p). DENYWPROF produces a lower FSS than any of the individual denial experiments after CI (Fig. 3.6), indicating that the RWPs lead to the largest improvements compared with the rest of the PECAN dataset.

Assimilating the rawinsonde observations collected during PECAN produces a large improvement similar to that from the RWPs. DENYSONDE simulates lower probabilities for both NCI and SCI compared to ALL at 0300 UTC (Fig. 3.7b,r). Additionally, DENYSONDE degrades the forecast for SCI at 0400 UTC (Fig. 3.7s) as the southern event is almost entirely missed. As in DENYWPROF, DENYSONDE also produces a large drop in FSS (Fig. 3.6) shortly after CI. Although the FSS for DENYSONDE converges with ALL (Fig. 3.6) due to high ensemble probabilities (>90%) within the MCS, the extent of the MCS in DENYSONDE is still reduced compared to ALL. Therefore, the large improvements from assimilating rawinsonde observations are partially maintained throughout the later periods of the forecast.

Finally, assimilating the special PECAN surface observations has a small benefit similar to those that result from assimilating the Doppler lidars. At early lead times, surface observations have a small, detrimental impact as seen by a higher FSS in DENYSFC compared to ALL at 0215 UTC (Fig. 3.6). These impacts again result from differences in how the DENYSFC experiment resolves the ongoing surface-based convection. Beginning at 0300 UTC (Fig. 3.7b,v), the positive impacts from assimilating the PECAN surface observations are mainly confined to SCI, as DENYSFC shows similar probabilities to DENYLIDAR. Again, these impacts from assimilating surface observations are small after the convection grows upscale, as the FSS from DENYSFC and ALL converge shortly after 0400 UTC (Fig. 3.6).

3.5 Ingredients-based analysis of the observations impacts

An ingredients-based approach (e.g., Johns and Doswell 1992) is used to determine which convective components (lift, moisture, instability) most contribute to the forecast impacts discussed in the previous section. By performing such an analysis, we determine exactly why certain observation types aid in the successful forecast of the 26 June nocturnal CI event.

3.5.1 Observation impacts on lifting mechanisms

We focus first on the lifting mechanisms responsible for the two individual CI clusters. Although the large-scale ascent mentioned in section 3.2 likely contributes to destabilization for parcels north of the synoptic boundary, additional mesoscale mechanisms are needed to lift the parcels to their LFC. For example, the observed sounding taken shortly before CI (Fig. 3.1c) shows that the most-unstable parcel, originating at 873 hPa, needed to be lifted to 762 hPa to reach its LFC ($\sim$1 km of lift). Analyses suggest that SCI forms along an outflow boundary produced by the surface-based convection along the synoptic front (as hypothesized by Trier et al. 2018; see Fig. 3.2c). NCI initiates shortly afterwards along the LLJ terminus (see Fig. 3.1b).

To first illustrate the large-scale, isentropic ascent, the heights of the 312 K virtual potential temperature (θ_v) isentrope are plotted (Fig. 3.8). In each simulation, the 312 K θ_v level is located at or just above the ground in east-central Kansas. The height of that same isentrope increases to the north as parcels are lifted isentropically above the synoptic boundary by the LLJ. Along the Kansas-Nebraska border and near the location of NCI, the 312 K isentrope is lifted to $\sim$1250 m AGL in each experiment, demonstrating little observation impacts on the larger-scale ascent. Fig. 3.8 also shows the horizontal mass convergence at 850 hPa over the LLJ terminus region for each data denial experiment. Again, no denial experiment has a large impact on the LLJ or the convergence located at its terminus. Although DENYWPROF and DENYSONDE simulate slightly weaker wind speeds just south of the jet terminus, those differences do not manifest in the convergence field. Because the ascent resulting from horizontal convergence is a function of the integrated convergence profile, vertical profiles are also shown in Figs. 3.9 (NCI) and 3.10 (SCI). Similar profiles of convergence are simulated by each experiment along the LLJ terminus and near NCI (Fig. 3.9b).

While the PECAN observations have little impact on forcing mechanisms for NCI, they have a larger impact on the convergence ahead of the outflow produced by the earlier surface-based convection (Figs. 3.10b, 3.11; see first column of Fig. 3.7 for the forecasts of this surface-based convection). The wind speeds and convergence ahead of the cold pool are maximized at 250 m AGL (Fig. 3.10) and are shown in Fig. 3.11. Compared to ALL and DENYAERI (Fig. 3.11a,b), wind speeds in the northern section of the cold pool are over 8 ms^{-1} slower in DENYLIDAR, DENYWPROF, and DENYSFC Fig. 3.11c-f). Due to the slower wind speeds south of this outflow boundary, the convergence profile along the boundary is weaker in those denial experiments (Fig. 3.10b). Although DENYSONDE simulates slower wind speeds within the cold pool compared to ALL (Fig. 3.11e), the experiment also enhances the winds ahead of the outflow boundary. Thus, DENYSONDE produces a similar magnitude of convergence as ALL for SCI (Fig. 3.11a,e, 3.10b). These convergence differences partially explain why assimilating the Doppler lidars, RWPs, and surface observations aid in enhancing SCI. Without the additional ascent along the outflow boundary, parcels in DENYLIDAR, DENYWPROF, and DENYSFC need additional time to reach their LFC.

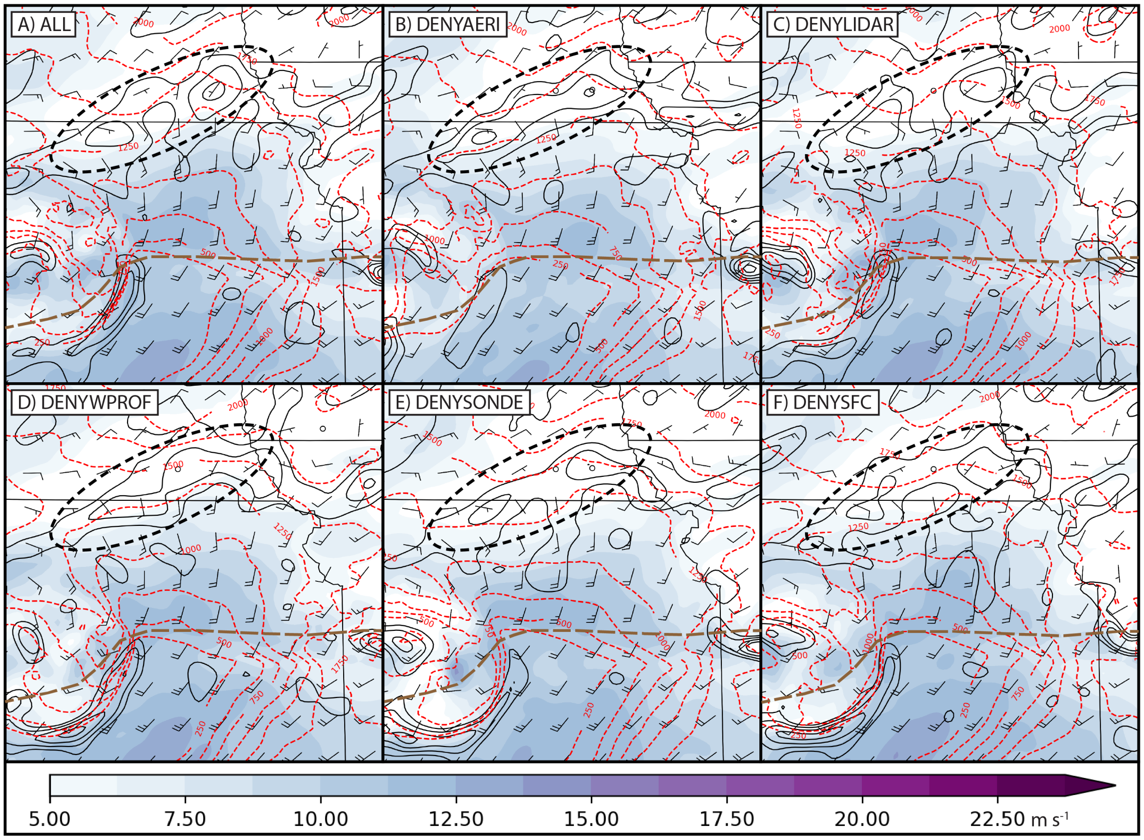


Figure 3.8: Ensemble mean forecasts of 850 hPa winds (fill; ms^{-1}) and convergence (contoured in black every +5-6 s^{-1}) for (a) ALL; (b) DENYAERI; (c) DENYLIDAR; (d) DENYWPROF; (e) DENYSONDE; and (f) DENYSFC valid at 0200 UTC 26 June. The plotting domain is shown by the gray box in Fig. 3.1b. The half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. Also overlaid in dashed red contours are the heights (every 250 m AGL) of the 312 K virtual potential temperature isentrope. The circled areas indicate the general location of the LLJ terminus corresponding to the CI event of interest, while the dashed brown lines indicate the approximate location of the synoptic boundary discussed throughout the text.

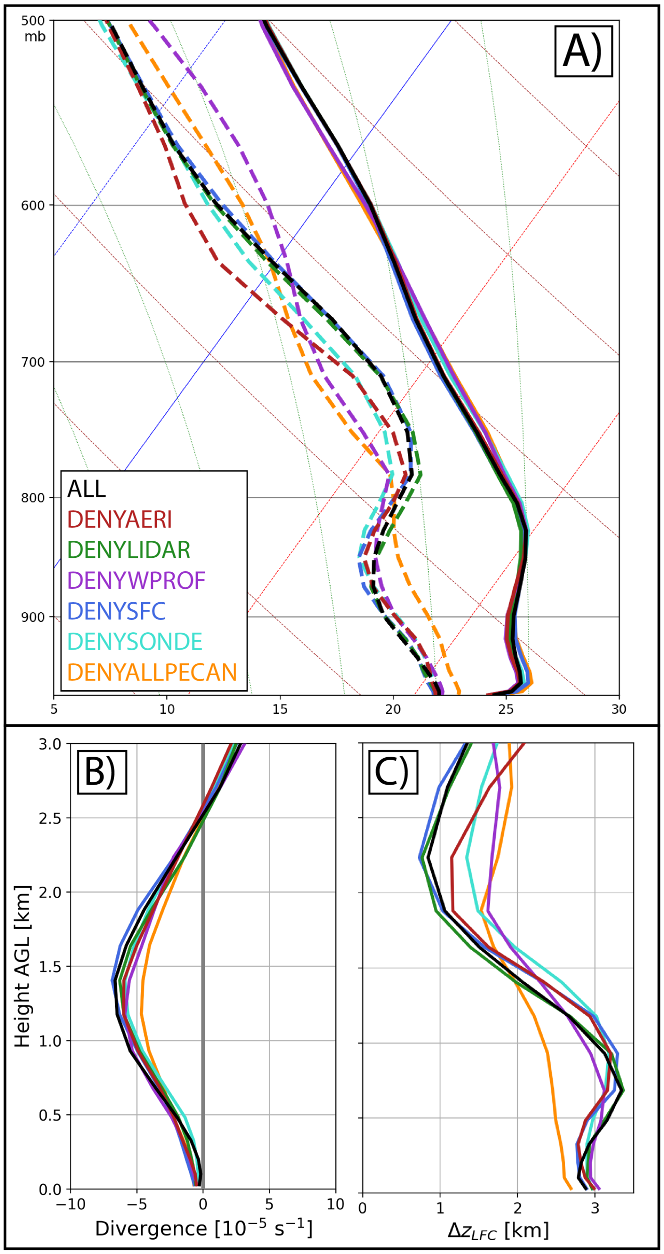


Figure 3.9: (a) Profiles of temperature (solid lines) and dewpoint temperature (dashed lines); (b) divergence ($10^{-5}\ s^{-1}$); and (c) Δz_{LFC} (km) for each data denial experiment valid at 0200 UTC 26 June. The ensemble mean profiles are averaged over the northern black box in Fig. 3.7c.

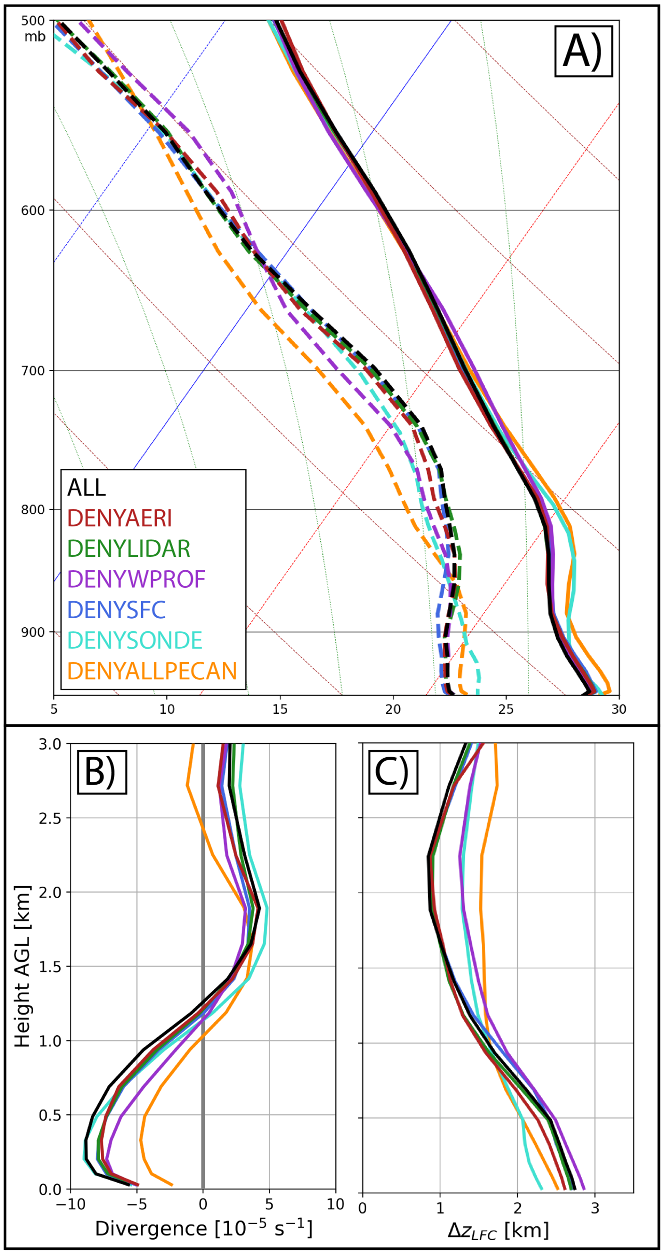


Figure 3.10: As in Fig. 3.9 but averaged over the southern black box in Fig. 3.7c and valid at 0130 UTC 26 June.

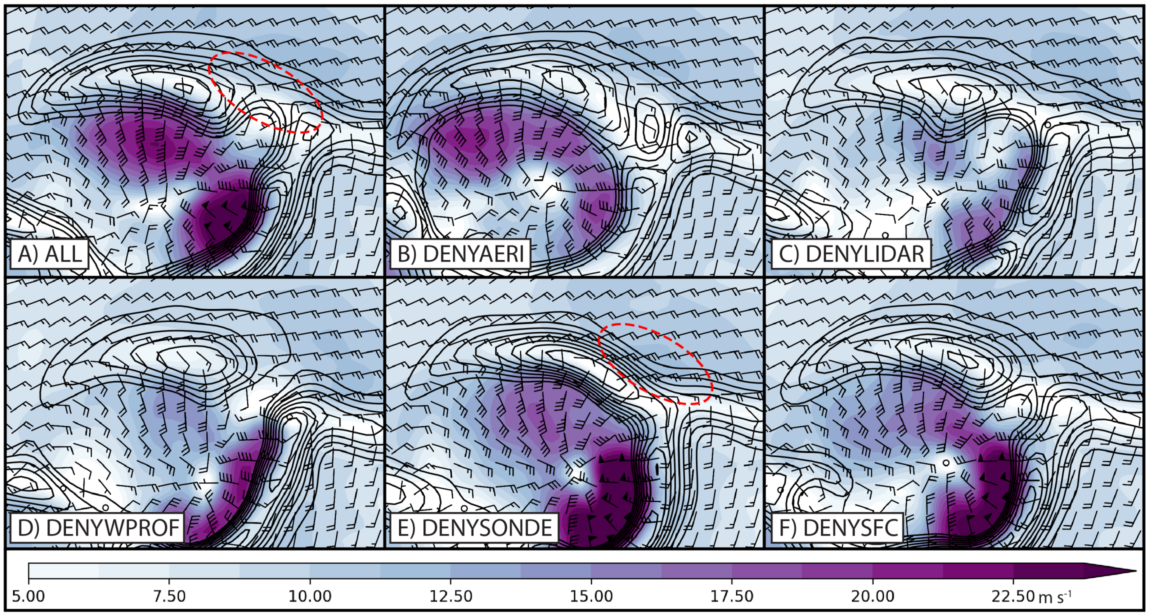


Figure 3.11: Ensemble mean forecasts of 250 m AGL winds (fill; ms^{-1}) and convergence (contoured in black every +5-6 s^{-1}) for (a) ALL; (b) DENYAERI; (c) DENYLIDAR; (d) DENYWPROF; (e) DENYSONDE; and (f) DENYSFC valid at 0130 UTC 26 June. The plotting domain is shown by the gray box in Fig. 3.2d. The half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. The plotting domain is zoomed into the outflow boundary produced by the ongoing surface-based cells discussed in the text. See text for a description of the circled areas.

3.5.2 Observation impacts on the thermodynamic environment

Although assimilating the PECAN observations has little impact on the convergence near LLJ terminus, the denial experiments show large sensitivities to the elevated moist layer in the same area (Figs. 3.9a,3.12). DENYAERI, DENYWPROF, and DENYSONDE, which all produce a large decrease in forecast skill for NCI, simulate drier midlevels near the LLJ terminus compared to ALL (Fig. 3.9a). The dry air in these three experiments leads to additional inhibition that needs to be eroded before parcels can reach their LFC. Another way of presenting the inhibition is through Δz_{LFC}, which describes the distance between a parcel's LFC and it's starting height. The Δz_{LFC} parameter can be interpreted as the amount of lifting needed for a parcel to produce an accelerating updraft. In ALL, parcels originating at 2.25 km AGL near NCI need to be lifted only 900 m to reach their LFC. This value corresponds well with the sounding in Fig. 3.1c. Without AERI, RWPs, or rawinsondes assimilated, these same parcels need to be lifted between 1200 and 1600 m (Fig. 3.9c). For SCI, only the assimilation of RWPs or rawinsondes significantly modifies the thermodynamic environment (Fig. 3.10a,c). While assimilating both RWPs and rawinsondes results in moister midlevels, the rawinsondes also strongly cool the layer between 900 and 800 hPa and thus further improve the environment for SCI. In ALL, parcels originating at 2 km AGL for SCI need 900 m of lift to reach their LFC, while the same parcels in DENYWPROF or DENYSONDE need 1300 m of lift (Fig. 3.10c).

The impact on the elevated moist layer is further shown by the plan view plot in Fig. 3.12. Compared to ALL, the Doppler lidar and surface observations have little impact on the midlevel dewpoint temperatures throughout northern Kansas. When denying the AERI, RWPs, or rawinsondes, however, the 700 hPa dewpoint temperatures are reduced by upwards of 6 °C in some locations. DENYWPROF and DENYSONDE both modify the elevated moist layer over a large region that corresponds to both NCI and SCI (red circles in Fig. 3.12d,e). Conversely, the observation impacts from assimilating AERI observations are

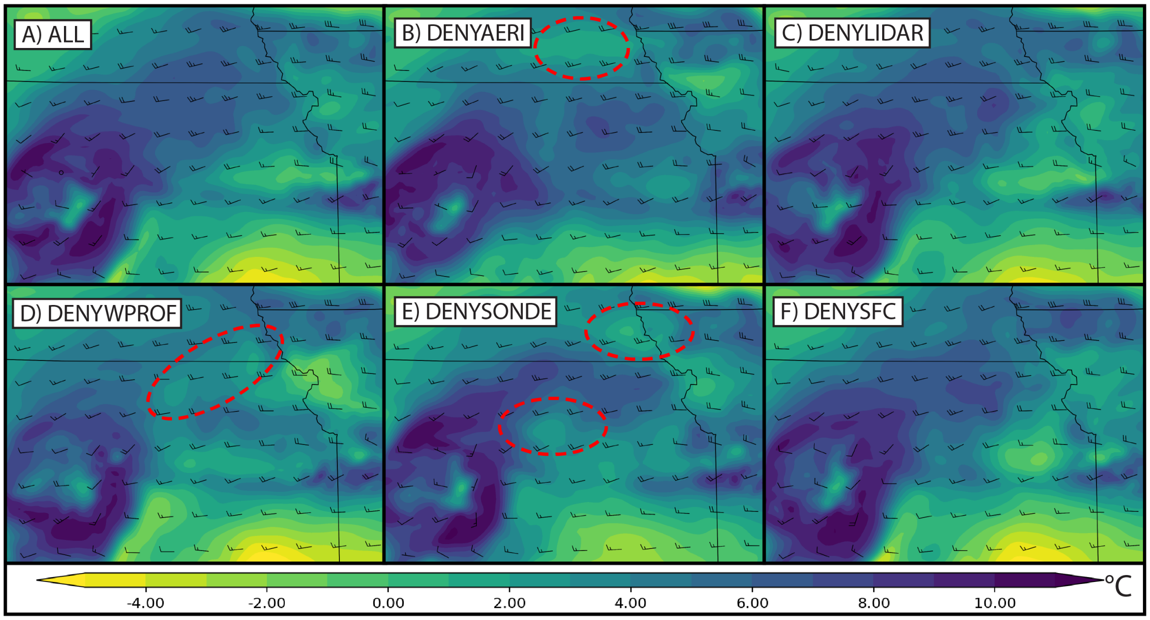


Figure 3.12: Ensemble mean forecasts of 700 hPa dewpoint temperature (°C) and wind barbs for (a) ALL; (b) DENYAERI; (c) DENYLIDAR; (d) DENYWPROF; (e) DENYSONDE; and (f) DENYSFC valid at 0200 UTC 26 June. The half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. See text for a description of the red circles.

mainly confined to the region near the LLJ terminus corresponding to NCI (red circle in Fig. 3.12b).

3.6 Analysis of observation impacts on the DA cycles

The previous section indicates that assimilating the PECAN observations enhances both the elevated moist layer (AERI, RWPs, and rawinsondes) and the convergence along an outflow region produced by earlier, surface-based convection (Doppler lidars, RWPs, and surface observations). These primary impacts likely lead to the improved forecast skill in ALL. However, it is not initially clear why, for example, assimilating the RWPs modifies the moisture field. Thus, this final section of results explores the observation impacts throughout the DA cycles to briefly explain how the assimilation of PECAN observations impacts these important fields.

3.6.1 AERIs

Differences between the 700 hPa water vapor mixing ratio analyses for ALL and DENYAERI are presented in Fig. 3.13a-d. In ALL, most of the additional moisture from assimilating AERI observations originates on the outer domain DA cycles (0300-2100 UTC 25 June). At 0600 UTC 25 June, ALL shows increased moisture above the synoptic boundary compared to DENYAERI. The additional moisture is maximized at 700 hPa (Fig. 3.13a). By 0900 UTC, the moisture differences along the boundary in ALL reach nearly +4 $g\ kg^{-1}$ (Fig. 3.13b). After 0900 UTC, the midlevel steering flow in ALL (see Fig. 3.12a) advects the additional moisture northeastward, eventually reaching the Kansas-Nebraska border by the final DA cycle on the outer domain (Fig. 3.13d). The region of additional moisture in ALL, compared to DENYAERI, correlates well with the location of NCI.

The additional moisture added above the synoptic boundary appears to be primarily related to negative increments in the midlevel moisture profile at FP2 (Fig. 3.14). First, the background ensemble in ALL indicates that the 750 hPa mixing ratio at the top of the AERI profiles from FP2 is negatively correlated with the elevated moist layer above the synoptic boundary (Fig. 3.14a; correlation calculated against 700 hPa mixing ratio where the impact is maximized). We hypothesize that this anti-correlation in the background is due to the ensemble indicating a strong moisture gradient along the frontal boundary, such that drier members at FP2 are moister above the front. Next, the midlevels in the background are also 10 °C too moist compared to corresponding AERI retrievals at FP2 (Fig. 3.14b). Therefore, while the AERI observations at FP2 aid in drying the midlevels in southwestern Kansas, the background covariance structure allows the same profiles to strongly moisten the midlevels above the synoptic boundary (Fig. 3.14a). This finding illustrates the primary advantage of using an ensemble-based DA method like the EnKF, as it can generate flow-dependent background error covariances.

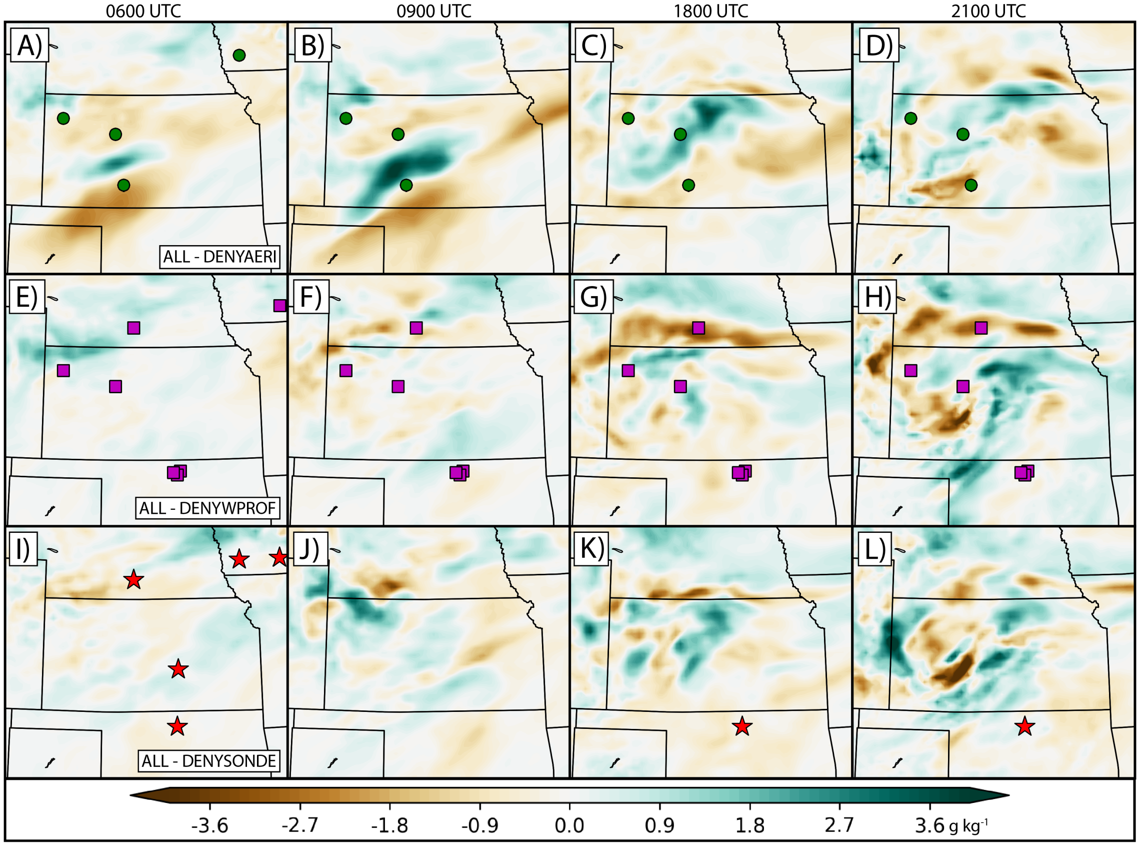


Figure 3.13: Differences in analyzed ensemble mean water vapor mixing ratio ($g\ kg^{-1}$) between ALL and (a-d) DENYAERI at 700 hPa; (e-f) DENYWPROF at 750 hPa; and (i-l) DENYSONDE at 750 hPa. The plots are valid at (a,e,i) 0600 UTC; (b,f,j) 0900 UTC; (c,g,k) 1800 UTC; and (d,h,l) 2100 UTC 25 June. Also overlaid are the locations of the AERIs (a-d; green dots), RWPs (e-f; magenta squares), and PECAN rawinsondes (i-l; red stars) assimilated in ALL.

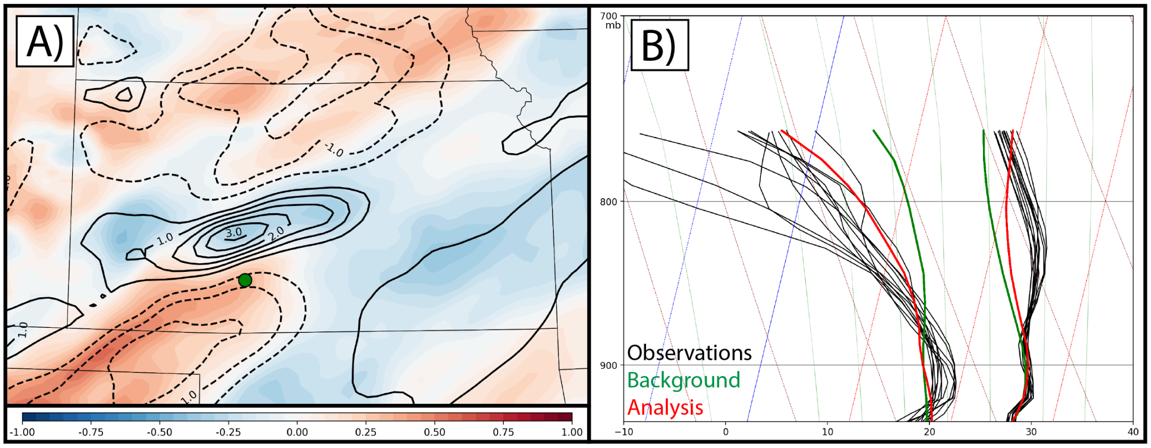


Figure 3.14: (a) Background ensemble correlations between 750 hPa water vapor mixing ratio at FP2 (green dot) and 700 hPa water vapor mixing ratio across the domain; and (b) examples of all AERI profiles assimilated from FP2 (black lines) at 0600 UTC 25 June. The green lines indicate the ensemble mean of the background, while the red lines indicate the ensemble mean of the analysis. The rightmost cluster of lines in (b) represents the temperature (°C), while the leftmost cluster of lines represents the dewpoint temperature (°C). Also overlaid in (a) are the ensemble mean analysis increments (background minus analysis; black contours) for water vapor mixing ratio ($g\ kg^{-1}$). The dashed contours indicate negative increments and the solid contours indicate positive increments. Both plots are taken from the ALL experiment.

3.6.2 RWPs and rawinsondes

As discussed earlier, assimilating the RWPs in ALL results in additional moisture at 750 hPa throughout a large region of northeastern Kansas (Fig. 3.13e-h). This moisture primarily manifests during the final outer domain DA cycles between 1800-2100 UTC 25 June (Fig. 3.13h). However, unlike the assimilation of AERI observations which directly add moisture, the additional moisture from assimilating RWPs results from enhancements to the moisture advection field (Fig. 3.15). In northwestern Kansas, the 1800 UTC wind profiler observations at FP3 and FP5 produce innovations of +2-4 ms^{-1} in the zonal wind, which in turn, lead to a large, positive increment in ALL (Fig. 3.15a). Because most of the midlevel moisture is also located in northwestern Kansas (Fig. 3.15), the enhancement of the zonal wind in ALL increases the moisture advection into central and eastern Kansas. Without the assimilation of the RWPs in DENYWPROF (Fig. 3.15b), the zonal wind decreases during the 1800 UTC cycle throughout much of western Kansas. Thus, DENYWPROF simulates weaker moisture advection that eventually leads to the large differences in the 750 hPa moisture field at 2100 UTC (Fig. 3.13h).

As with the impact from assimilating RWPs, the assimilation of rawinsonde data also leads to modifications to the midlevel zonal wind fields during the 1800 UTC cycle (Figs. 3.13k,3.15c). However, only one rawinsonde was launched during this cycle while many wind profiles were collected throughout the domain (see Fig. 3.2b). The single rawinsonde assimilated at 1800 UTC from FP1 shows a large, negative innovation ($\sim$-3 ms^{-1}) between the observed and simulated zonal wind at 700 hPa (Fig. 3.16b). Because both the innovation at FP1 and correlations with the 750 hPa wind in western Kansas are negative (Fig. 3.16a), we deduce that assimilating the FP1 rawinsonde at 1800 UTC is at least partially responsible for the positive increment in the zonal wind shown in Fig. 3.15a. Therefore, in DENYSONDE, a negative increment in the zonal wind occurs in southwestern Kansas at 1800 UTC (Fig. 3.15c) that weakens the moisture advection into central Kansas. As in

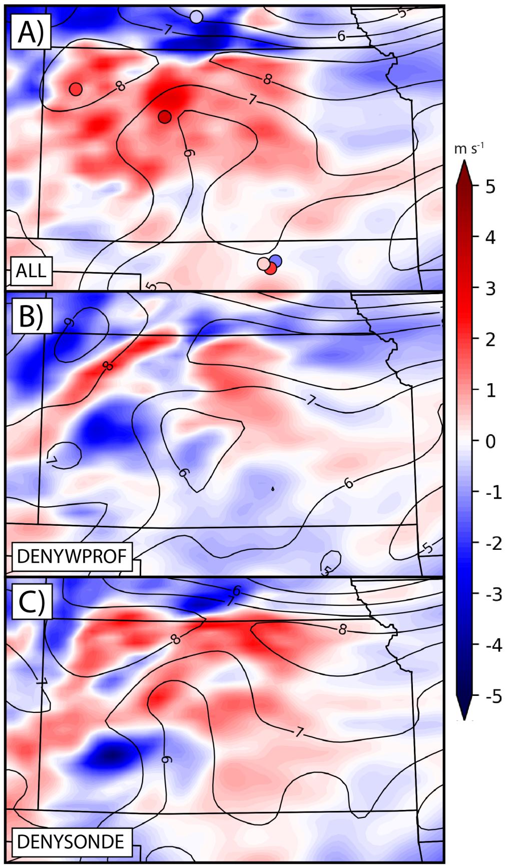


Figure 3.15: Ensemble mean analysis increments (analysis minus background) of the 750 hPa zonal wind (ms^{-1}) for (a) ALL; (b) DENYWPROF; and (c) DENYSONDE valid at 1800 UTC 25 June. The black contours indicate the ensemble mean 750 hPa water vapor mixing ratio ($g\ kg^{-1}$) for each respective experiment. Also overlaid in (a) are the ensemble mean innovation values (observation-background; dots) for the RWP observations closest to 750 hPa.

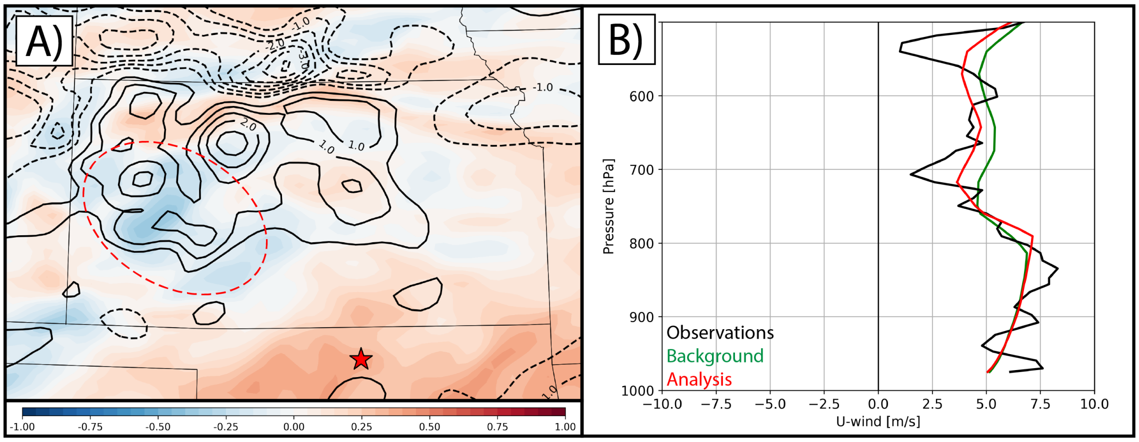


Figure 3.16: (a) Background ensemble correlations between 700 hPa zonal wind at FP1 (red star) and 750 hPa zonal wind across the domain; and (b) the zonal wind profile assimilated from the FP1 rawinsonde at 1800 UTC 25 June. The green lines indicate the ensemble mean of the background, while the red lines indicate the ensemble mean of the analysis. Also overlaid in (a) are the ensemble mean analysis increments (background minus analysis; black contours) for zonal wind (ms^{-1}). The dashed contours indicate negative increments and the solid contours indicate positive increments. Both plots are taken from the ALL experiment.

DENYWPROF, the weaker moisture advection then leads to reduced midlevel moisture during the later DA cycles in DENYSONDE compared to ALL (Fig. 3.13k,l).

3.6.3 Doppler lidars and surface observations

To determine why assimilating Doppler lidars or surface observations enhances the wind speeds within the outflow boundary, we analyze the common elements that contribute to stronger cold pools. We find little sensitivity to either the precipitation within the surface-based cells or the relative humidity profile below cloud base (not shown). Instead, when these observation types are assimilated near the ongoing surface-based convection, convective-scale regions along the borders of the storms are moistened by the final assimilation cycle (Fig. 3.17b,h). This moisture impact is maximized at 600 hPa (Fig. 3.18). Assimilating RWPs and rawinsonde observations produces similar effects near the ongoing convection (Fig. 3.17d,f), though additional moisture is already present due to the effects discussed previously. The additional moisture from assimilating Doppler lidar and surface observations does not exist prior to the development of the surface-based convection (Fig. 3.17a,g), indicating that the impacts are related to convective-scale DA. This impact extends throughout much of the mid-troposphere, with DENYLIDAR and DENYSFC simulating decreased dewpoint temperatures by an average of 2-4 °C between 500 and 800 hPa (Fig. 3.18). We hypothesize that the additional, convective-scale moisture added by these observations enhances the ongoing surface-based convection and later produces the stronger outflow seen only in ALL and DENYAERI (Fig. 3.9). Additionally, this increased moisture would likely reduce the impact of entrainment effects that could act to dissipate the ongoing convection.

3.7 Discussion

By assimilating remote sensing profilers, high-frequency rawinsondes, and surface observations collected on 26 June, we find large improvements over a baseline experiment in terms of location, orientation, and timing of a nocturnal CI forecast. The most skillful forecast occurs when assimilating every PECAN dataset used in this chapter, thus indicating that each observation type plays a positive role in improving the CI forecast. Our results also suggest that the linear CI episode was initiated by two separate forcing mechanisms. NCI was initiated largely by the LLJ, while SCI formed along an outflow boundary produced by earlier, surface-based convection.

We conduct experiments within a data denial framework to evaluate the relative impact of assimilating each PECAN observation type within the full dataset. Assimilating AERI, RWP, and rawinsonde data produces the largest and most sustained impact by enhancing the elevated moist layer in the region of CI. The RWPs and rawinsondes affect both NCI and SCI by strengthening the moisture advection across northern Kansas. Assimilating AERI observations directly adds moisture above the synoptic boundary that is then advected into the NCI. This study is among the first to assimilate real AERI observations and demonstrates that high-frequency profiles of temperature and water vapor can improve short-term forecasts

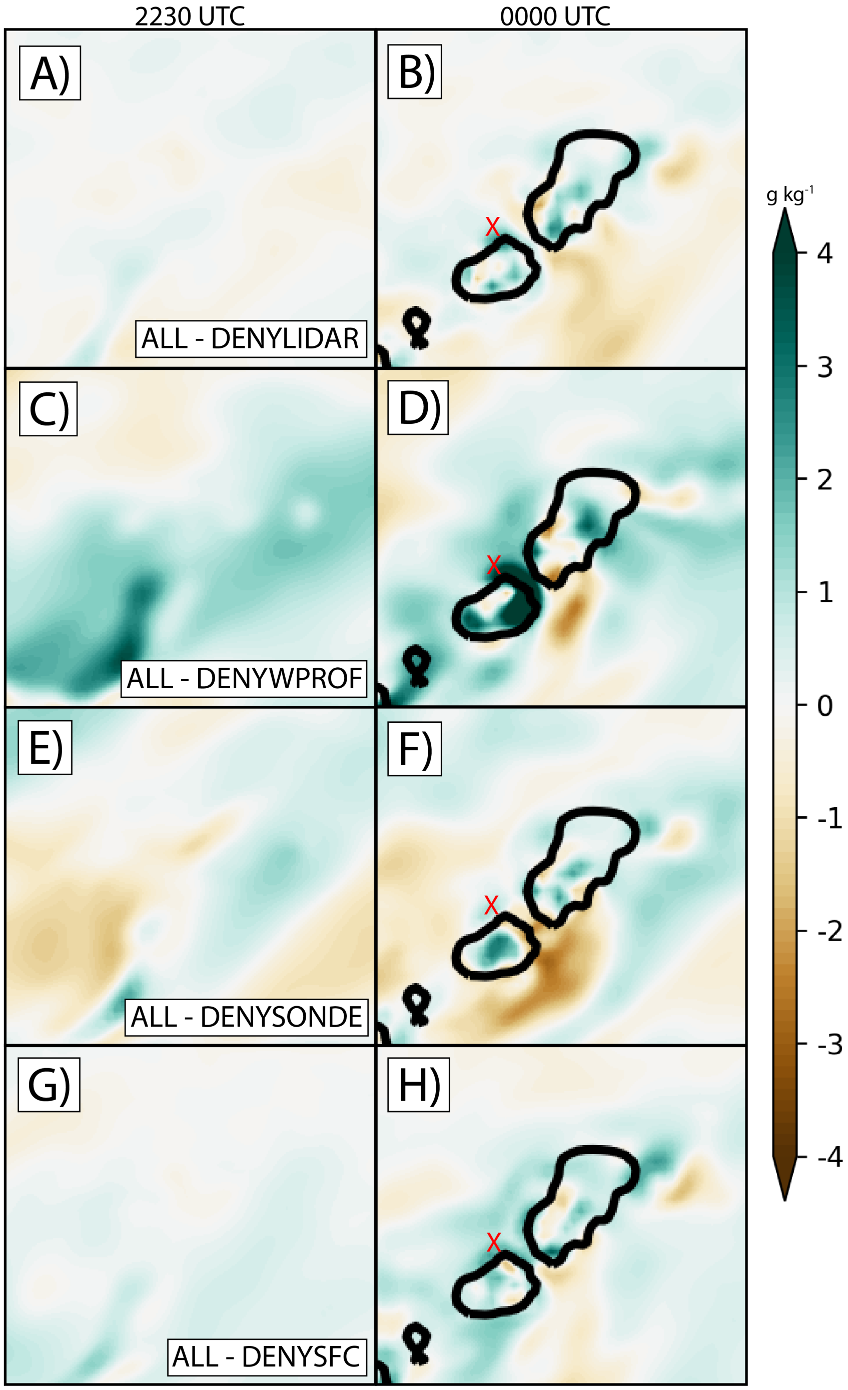


Figure 3.17: As in Fig. 3.13 but for differences in ensemble mean water vapor mixing ratio ($g\ kg^{-1}$) at 650 hPa between ALL and (a,b) DENYLIDAR; (c,d) DENYWPROF; (e,f) DENYSONDE; and (g,h) DENYSFC. The plots are valid at (a,c,e,g) 2230 UTC 25 June; and (b,d,f,h) 0000 UTC 26 June. Also overlaid are 15 dBZ contours of composite reflectivity from the ensemble mean of ALL. The red crosses in (b,d,f,h) indicate the location of the profile in Fig. 3.18.

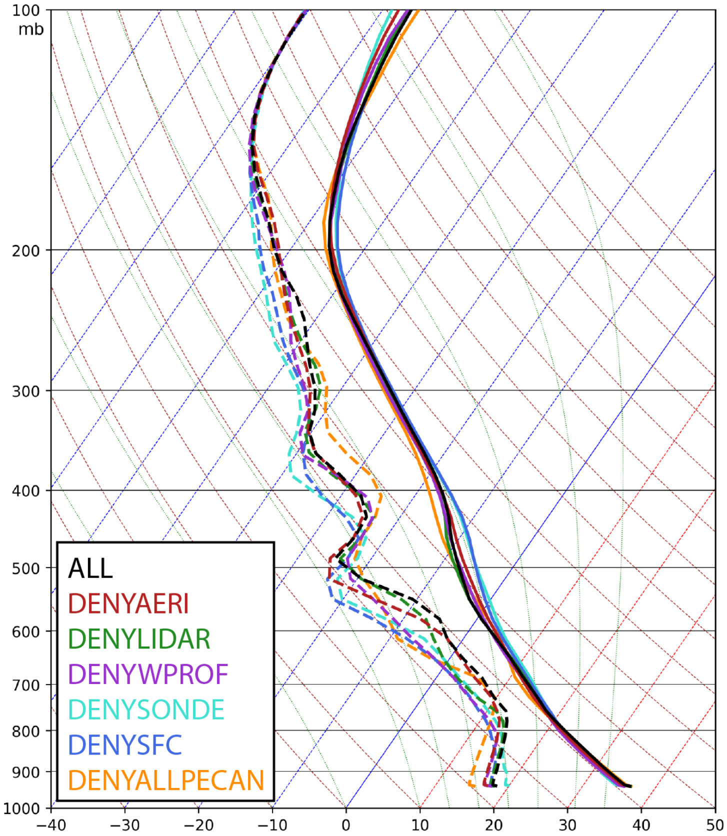


Figure 3.18: Ensemble mean profiles of temperature (solid lines) and dewpoint temperature (dashed lines) taken from the red crosses in Fig. 3.17 and valid at 0000 UTC 26 June. The sounding is averaged over a neighborhood with a radius of two grid points (8 km).

of convection. Additionally, the special rawinsondes assimilated here were launched more frequently and at non-standard times compared to the operational network, thus providing further evidence for the value of assimilating high-frequency profiles.

The largest improvements result during DA cycling on the outer, mesoscale domain, indicating that assimilating profiler data can lead to forecast improvements even when not assimilating the data on a convection-permitting grid. However, additional improvements are found when assimilating the PECAN data at 4 km. When assimilating surface and Doppler lidar observations, the pre-existing, surface-based convection produces a stronger outflow that enhances the ascent for the SCI. We hypothesize that the enhanced outflow is related to increased moisture near the analyzed convection that then enhances the ongoing storms during DA on the inner domain. Similar enhancements are also seen when assimilating the RWP observations. However, the improvements from assimilating surface and Doppler lidar observations diminish after the two simulated CI clusters merge into a larger MCS.

Still, various aspects of the results should be further explored. First, the location of each observation likely plays an important role on its impact. For example, the RWPs assimilated here are possibly more impactful than Doppler lidars due to the additional RWP at FP4 (far northwestern Kansas site in Fig. 3.15a). An additional Doppler lidar at the same location could allow for a similar increment for the zonal wind in northern Kansas. However, the higher maximum height of RWPs (upwards of 10 km AGL) also likely aids in the larger impact compared to Doppler lidars. Next, while we find an enhanced outflow boundary when assimilating Doppler lidar and surface observations, the impacts of convective-scale DA near ongoing convection is an area of research that has yet to be fully explored. Ensemble correlations near ongoing thunderstorms could be considered spurious due to the chaotic nature of convection. Thus, the impacts of assimilating the PECAN observations on the strength of the outflow boundary should be further studied.

For similar cases that show large thermodynamic errors, we expect that assimilating profiler observations can lead to improvements for short-term forecasts of CI. However, the strong forcing mechanisms for this event are well-captured by each experiment, such that only the thermodynamic enhancements are needed for a successful CI forecast. It is unclear whether assimilating such data could improve convergence mechanisms for other CI events, or if the observation impacts would be as large when the mechanisms are not well captured. As nocturnal convection can be initiated by many other features such as atmospheric bores or internal gravity waves, the next chapter of this dissertation conducts a systematic evaluation of the impact of assimilating PECAN field observations on forecasts of nocturnal CI.

Chapter 4

Systematic evaluation of the impact of assimilating a network of ground-based remote sensing profilers for forecasts of nocturnal convection initiation during PECAN

4.1 Introduction

While conventional rawinsondes provide data only a few times per day, convective-scale environments often evolve on the order of minutes to hours (Orlanski 1975). To alleviate this data gap, the meteorological community has recently pushed to expand current observing capabilities into a Nationwide Network of Networks (National Research Council 2009; Stalker et al. 2013). Part of this proposal involves the introduction of ground-based remote sensing profilers to provide high-frequency observations of the lower troposphere. Although their errors are often larger than corresponding in-situ measurements, remote sensing instruments can provide thermodynamic and kinematic data multiple times per hour. Additionally, as various studies show that short-term forecasts of convection are highly sensitive to small changes in the mesoscale environment (e.g., Crook 1990; Martin and Xue 2006), assimilating such instruments can potentially have large benefits for convective-scale NWP.

Recently, the PECAN experiment (Geerts et al. 2017) employed AERIs and vertical wind profilers to study nocturnal convection in the Great Plains of the United States. As discussed at length throughout this dissertation, convection-allowing NWP systems often struggle to predict such events due to model errors related to important nocturnal mechanisms. However, a lack of observations that sample the mechanisms and environments responsible for nocturnal convection also contribute to reduced forecast skill. In the Great Plains, nocturnal CI most commonly occurs between 0400 and 0900 UTC (Reif and Bluestein

2017), while conventional rawinsondes are typically only launched at 0000 and 1200 UTC. Moreover, most nocturnal CI events are driven by elevated ascent mechanisms such as the LLJ or undular bores (Wilson and Roberts 2006; Weckwerth et al. 2019). Such features cannot be easily observed by conventional surface data, and various studies note the necessity of lower-troposphere profiling observations to improve this data gap (e.g., (Wilson and Roberts 2006; Weckwerth and Parsons 2006; Keclik et al. 2017). Given these deficiencies, recent works show that assimilating high-frequency, remote sensing datasets can be particularly useful for improving forecasts of nocturnal convection. For example, we show in Chapter 3 that assimilating both AERIs and kinematic profilers can enhance moisture advection and support a successful CI forecast. Additionally, Chipilski et al. (2020) find that assimilating kinematic profilers improves analyses of a nocturnal LLJ and explicit forecasts of an undular bore, such that the resulting convective forecast is also improved. Additional studies assimilating AERIs and kinematic profilers include Hu et al. (2019) and Coniglio et al. (2019) who find improvements to short-term forecasts of tornadic supercells and convective evolution, respectively.

While previous works show the benefits of assimilating ground-based remote sensing datasets for single cases, many research areas remain unexplored. For example, nocturnal convection can be generated by a variety of ascent mechanisms at different scales (Reif and Bluestein 2017; Stelten and Gallus 2017; Weckwerth et al. 2019) such that the impact of assimilating these instruments is likely case- and environment-dependent. While Chapter 3 shows the benefits of assimilating the data for a single, large-scale frontal overrunning event, it remains unclear whether the impacts will be as large for smaller-scale events such as those initiated by gravity waves or outflow boundaries. To address this question, we expand on previous work by using carefully designed experiments to evaluate the systematic impact of assimilating a network of high-frequency (observations available every $\sim$5-30 min) boundary layer profilers. We explore these impacts using cases observed during the PECAN field experiment, as the variety of events provides a sufficient sample of the different convective modes and mesoscale environments responsible for nocturnal CI. Furthermore, Stalker et al. (2013) propose implementing new networks alongside existing infrastructure such as the operational rawinsonde network. Thus, this dissertation chapter specifically evaluates the impact of assimilating collocated rawinsondes, AERIs, and kinematic profilers. We also explore whether the impact of assimilating these data primarily result from one instrument type, or if both thermodynamic and kinematic profilers are necessary to improve forecasts of nocturnal CI.

Details on the sensitivity experiments are presented in section 4.2. Section 4.3 discusses verification methods we use to evaluate the impact of assimilating these data. The systematic results for assimilating the remote sensing profilers on forecasts of nocturnal CI are found in section 4.4. To better understand where the impacts of assimilating these data originate, we conduct a pre-convective analysis in section 4.5 and example cases are presented in section 4.6. A final discussion is found in section 4.7.

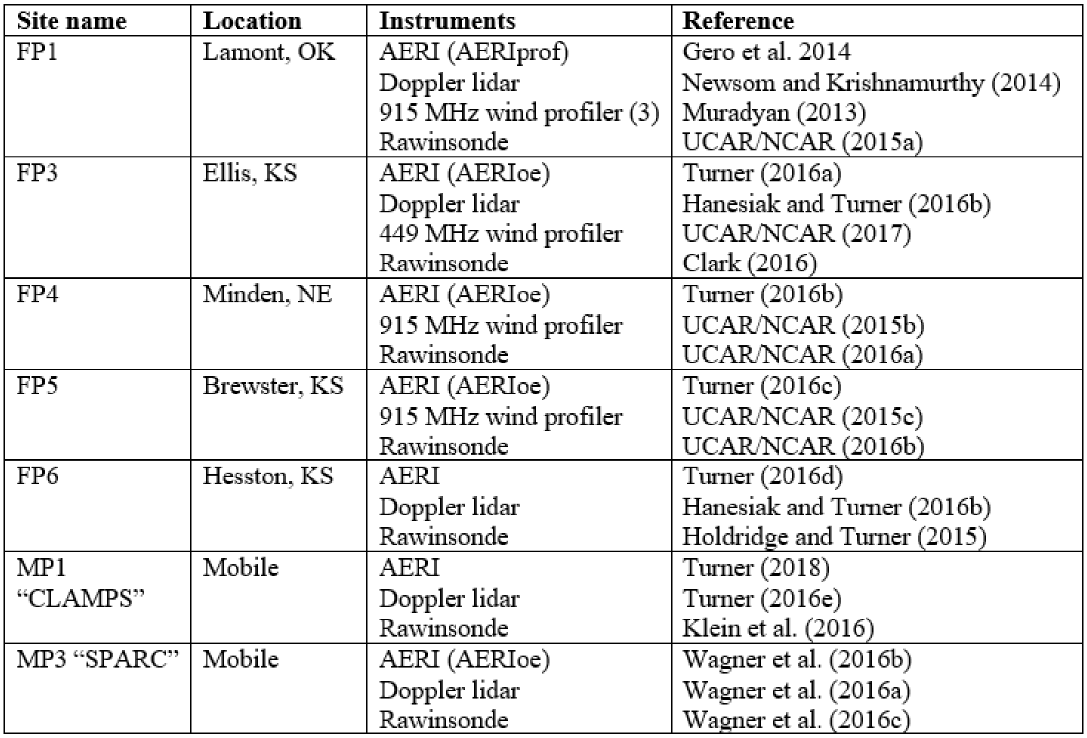
4.2 Observation pre-processing and sensitivity experiments

4.2.1 Observation pre-processing

The PECAN instruments assimilated in this chapter include rawinsondes, AERIs, Doppler lidars, and RWPs collected from both FP and MP sites (Table 4.1 and Fig. 4.1). Unlike the previous section, we no longer assimilate mobile surface observations given that this chapter aims to evaluate the impact of assimilating remote sensing instruments only. Other than this, each observation is assimilated using the same pre-processing methods described at length in Chapter 2.3 with a few differences described below.

First, as the goal of this chapter is to evaluate the impact of assimilating a collocated network of profiling instruments, we only assimilate data from sites that include a rawinsonde and each of the remote sensing profilers. However, only two PECAN sites feature each instrument (Table 4.1). To increase the potential number of sites, we treat Doppler lidars and RWPs as a single dataset that is henceforth referred to as “composite kinematic profilers”.

Table 4.1: As in Table 2.3 but only for the collocated observations assimilated in Chapter 4.



Site name	Location	Instruments	Reference
FP1	Lamont, OK	AERI (AERIprof) Doppler lidar 915 MHz wind profiler (3) Rawinsonde	Gero et al. 2014 Newsom and Krishnamurthy (2014) Muradyan (2013) UCAR/NCAR (2015a)
FP3	Ellis, KS	AERI (AERIoe) Doppler lidar 449 MHz wind profiler Rawinsonde	Turner (2016a) Hanesiak and Turner (2016b) UCAR/NCAR (2017) Clark (2016)
FP4	Minden, NE	AERI (AERIoe) 915 MHz wind profiler Rawinsonde	Turner (2016b) UCAR/NCAR (2015b) UCAR/NCAR (2016a)
FP5	Brewster, KS	AERI (AERIoe) 915 MHz wind profiler Rawinsonde	Turner (2016c) UCAR/NCAR (2015c) UCAR/NCAR (2016b)
FP6	Hesston, KS	AERI Doppler lidar Rawinsonde	Turner (2016d) Hanesiak and Turner (2016b) Holdridge and Turner (2015)
MP1 “CLAMPS”	Mobile	AERI Doppler lidar Rawinsonde	Turner (2018) Turner (2016e) Klein et al. (2016)
MP3 “SPARC”	Mobile	AERI (AERIoe) Doppler lidar Rawinsonde	Wagner et al. (2016b) Wagner et al. (2016a) Wagner et al. (2016c)

The observations for sites that feature both sets of wind profilers are combined using three steps. First, the high-frequency Doppler lidar profiles are temporally averaged using the same averaging window as the collocated RWP. Then, for any vertical levels where the two wind profiles overlap, we interpolate the denser profile (usually the Doppler lidar) onto the coarser observation and the two are averaged together. The observation error variances are also averaged throughout these layers to produce the error profiles for the composite dataset. Lastly, observations from where the two datasets do not overlap (above or below the averaging layer) are added back into the composite kinematic profile. This combination increases the number of available assimilation sites to seven (five fixed and two mobile sites; Table 4.1 and Fig. 4.1).

Additionally, we now assimilate the AERI from the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP; Sisterson et al. 2016) station located near Lamont, OK (known as FP1 during PECAN). This instrument operated continuously throughout PECAN

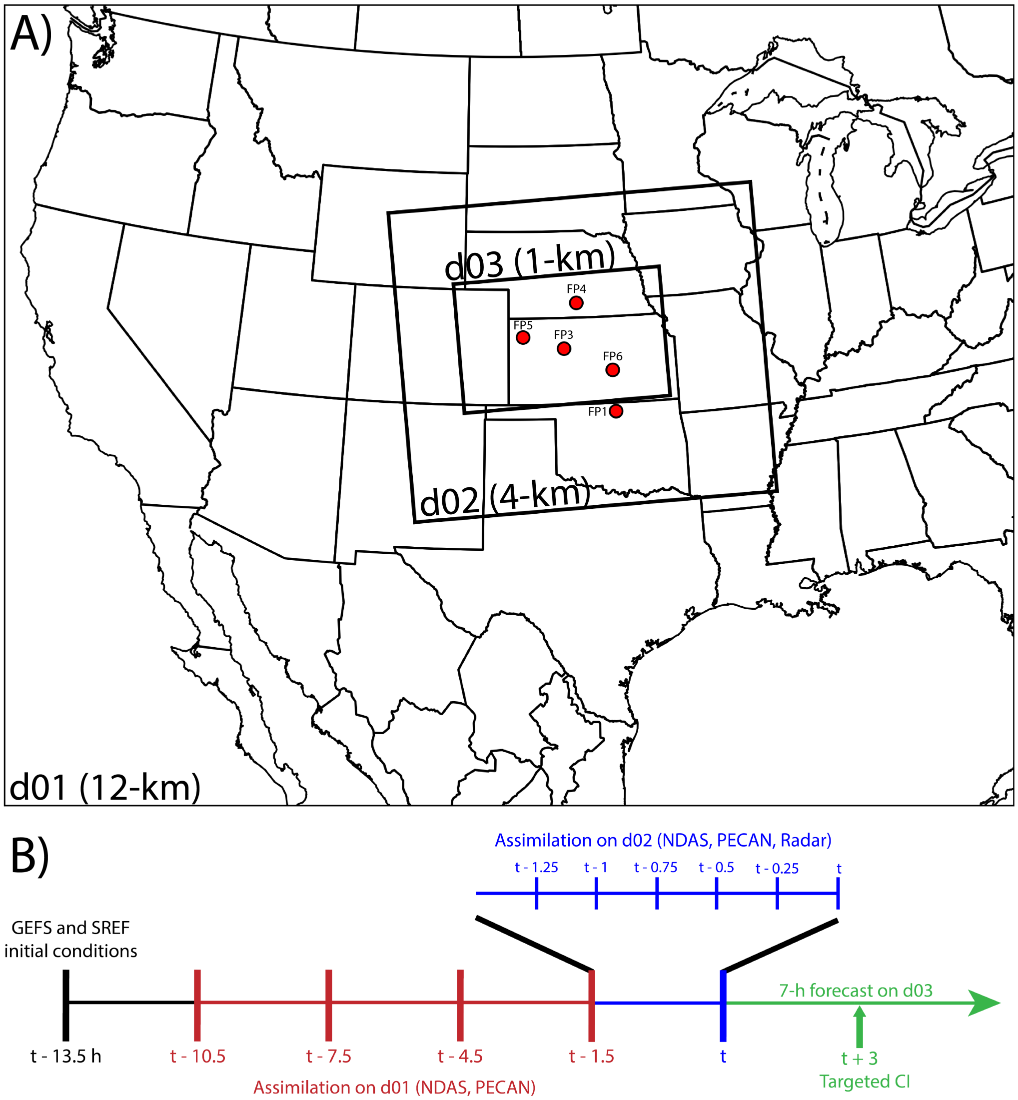


Figure 4.1: Overview of experimental design for Chapter 4 including (a) domain configuration for the outer (d01, 12-km) and intermediate (d02, 4-km) assimilation domains, as well as an example forecast domain (d03, 1-km) used for the 30 June nocturnal CI case. The forecast domain is the same size for other cases, though its exact location within d02 shifts depending on the location of the primary CI event. The location of each fixed PECAN observing site assimilated here are also overlaid. We note that two additional mobile PECAN sites are also assimilated for some cases, though their location varied by event. (b) Flowchart for the cycled assimilation, including the four 3-h assimilation cycles on d01, the six 15-min assimilation cycles on d02, and the 7-h forecast period. The primary CI event for each case (Table 4.2) occurs 3 h after the forecast initialization period.

and provided data for nearly every case examined here. We note that the AERI observations at FP1 are retrieved using the AERIprof retrieval algorithm (Smith et al. 1999; also called AERIplus in Feltz et al. 2003) instead of AERIoe. AERIprof differs from AERIoe by using a first guess from Rapid Refresh analyses (RAP; Benjamin et al. 2016) instead of a sounding climatology, having worse performance under low- or midlevel clouds, collecting less-frequent observations (15-min as opposed to 5-min retrievals), and not producing unique observation error profiles for each observing time. To assign the observation errors for AERI profiles collected from FP1, we instead use root-mean-square differences (RMSD) between the AERI and collocated soundings (given by Fig. 5 in Feltz et al. 2003). These errors are further inflated using the same method applied to the AERIoe retrievals (see section 3.3.1).

4.2.2 Case selection

Weckwerth et al. (2019) describe the variety of physical processes associated with nocturnal CI events in the Great Plains of the United States including frontal overrunning, events forming near a parent MCS, bores or density currents, or pristine CI events. An overview of each category is described here, and example cases simulated here are shown in Fig. 4.2 for each category. Frontal overrunning events occur when southerly flow advects warm, moist air above a frontal boundary, typically resulting in large-scale CI north of the boundary (Fig. 4.2a). Near MCS events form in proximity of an MCS but not along a radar fine line and can include “bow and arrow” (Keene and Schumacher 2013; example in Fig. 4.2b) and “T-initiation” (Coniglio et al. 2011) events. Bore or density current events are also associated with a parent MCS but instead develop along a radar fine line associated with an outflow boundary (Fig. 4.2c). Finally, pristine CI events include convective episodes not influenced by parent convective systems or obvious frontal boundaries (Fig. 4.2d). Weckwerth et al. (2019) provides further details and conceptual diagrams for each type of nocturnal CI.

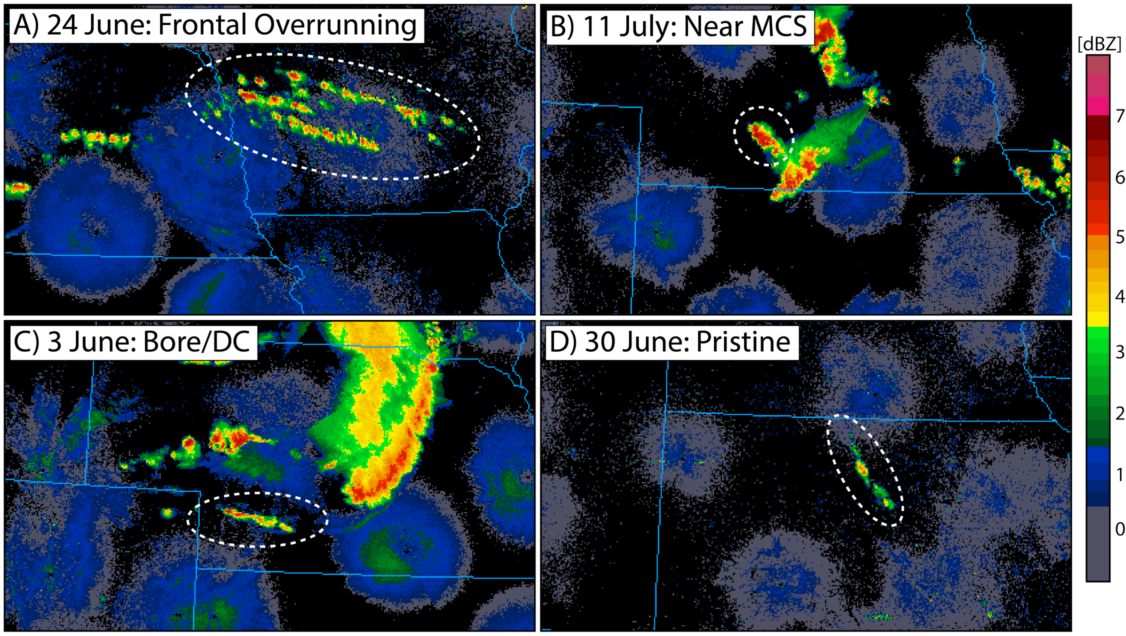


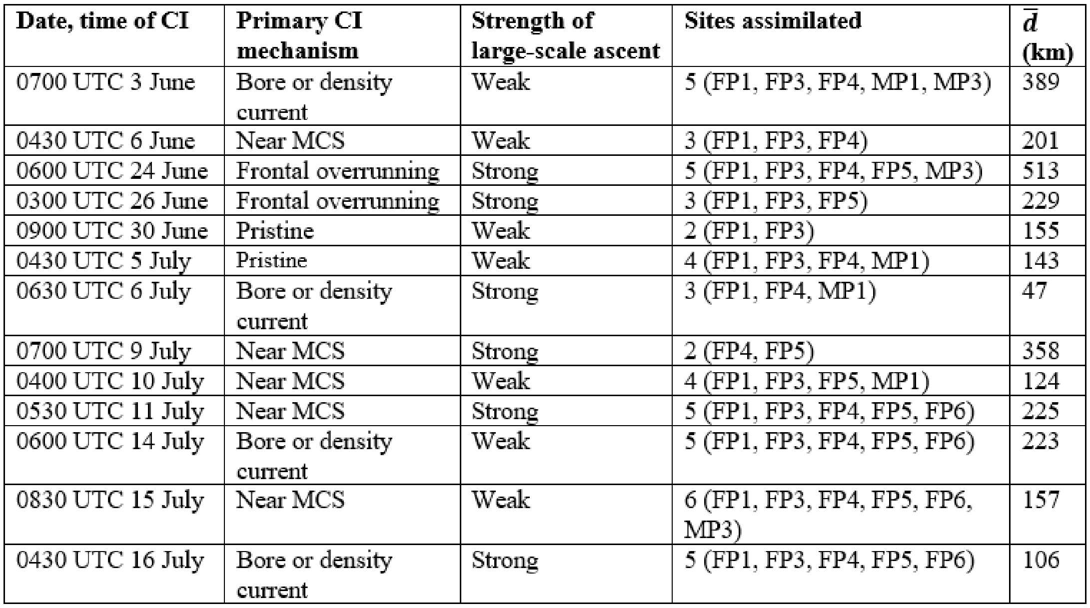
Figure 4.2: Composite reflectivity (dBZ) for example nocturnal CI events from each category including (a) frontal overrunning; (b) near MCS (bow and arrow event shown); (c) bore or density current; and (d) pristine. The dashed white oval regions indicate the primary nocturnal CI events that are likely initiated by the indicated ascent mechanisms. These images are courtesy of the image archive (available online at https://www2.mmm.ucar.edu/imagearchive/) maintained by the Mesoscale and Microscale Meteorology Division (MMM) of NCAR.

Given this large variety of processes responsible for nocturnal CI, the skill of a convection-allowing NWP system to predict these events likely varies significantly by case. We thus explore the impact of assimilating the thermodynamic and kinematic profilers through systematic experiments that include many nocturnal CI events observed during PECAN. From the list of nocturnal CI events observed during PECAN and compiled by Weckwerth et al. (2019), we select a case for evaluation if it features (1) at least two observing sites with a rawinsonde launched between 2330 UTC and 0030 UTC, and (2) at least one observing site that is no more than 300 km away from the center of the nocturnal CI event. These conditions ensure that we can evaluate the impact of assimilating a collocated network of remote sensing profilers alongside a representation of the current operational rawinsonde network and that the observations are close enough to impact the primary CI event. We note that while the cases are selected based on the presence of this primary CI event, we verify all observed CI events that occur within the forecast domain. After applying these conditions, we select 13 nocturnal CI cases for evaluation (Table 4.2). On average, the geographic center of the observing sites is located 221 km (σ = 124.1 km) away from primary nocturnal CI event. While this distance is larger than other studies that assimilate field campaign data (e.g., Chipilski et al. 2020), the cycled data assimilation allows for observation impacts to propagate far from the observing site, thus justifying the verification of all CI events within the forecast domain.

4.2.3 Experimental design

To determine the systematic impact of assimilating this collocated network of thermodynamic and kinematic profilers, we carefully design a set of data addition experiments (Table 4.3) based around a baseline experiment, SONDE that only features the assimilation of conventional observations and PECAN rawinsondes. We then separately assimilate collocated AERIs and composite kinematic profilers in the SONDE_TQPROF and SONDE_UVPROF experiments, respectively, by also assimilating the remote sensing data in addition to the

Table 4.2: List of nocturnal CI events simulated in Chapter 4 and their most likely ascent mechanisms from Weckwerth et al. (2019). The strength of the large-scale ascent for each case is also denoted. CI events that develop near a 500 hPa or 700 hPa upstream trough are classified as strongly-forced, while other events are classified as weakly-forced. The PECAN observing sites assimilated and their average geographic distance from the nocturnal CI event (km) are also listed for each case. Note that while these mechanisms correspond to the primary CI event used to select the cases, all CI objects occurring within the forecast period and domain are verified.



Date, time of CI	**Primary CI mechanism**	**Strength of large-scale ascent**	**Sites assimilated**	**$\overline{d}$ (km)**
0700 UTC 3 June	Bore or density current	Weak	5 (FP1, FP3, FP4, MP1, MP3)	389
0430 UTC 6 June	Near MCS	Weak	3 (FP1, FP3, FP4)	201
0600 UTC 24 June	Frontal overrunning	Strong	5 (FP1, FP3, FP4, FP5, MP3)	513
0300 UTC 26 June	Frontal overrunning	Strong	3 (FP1, FP3, FP5)	229
0900 UTC 30 June	Pristine	Weak	2 (FP1, FP3)	155
0430 UTC 5 July	Pristine	Weak	4 (FP1, FP3, FP4, MP1)	143
0630 UTC 6 July	Bore or density current	Strong	3 (FP1, FP4, MP1)	47
0700 UTC 9 July	Near MCS	Strong	2 (FP4, FP5)	358
0400 UTC 10 July	Near MCS	Weak	4 (FP1, FP3, FP5, MP1)	124
0530 UTC 11 July	Near MCS	Strong	5 (FP1, FP3, FP4, FP5, FP6)	225
0600 UTC 14 July	Bore or density current	Weak	5 (FP1, FP3, FP4, FP5, FP6)	223
0830 UTC 15 July	Near MCS	Weak	6 (FP1, FP3, FP4, FP5, FP6, MP3)	157
0430 UTC 16 July	Bore or density current	Strong	5 (FP1, FP3, FP4, FP5, FP6)	106

Table 4.3: List of experiments simulated in Chapter 4.

Experiment	Datasets assimilated
SONDE	PECAN rawinsondes launched between 2330 and 0030 UTC
SONDE_TQPROF	All observations from SONDE plus collocated thermodynamic profilers
SONDE_UVPROF	All observations from SONDE plus collocated kinematic profilers
SONDE_ALLPROF	All observations from SONDE_TQPROF and SONDE_UVPROF

baseline observations. To ensure that these data addition experiments demonstrate the impact of assimilating the remote sensing instruments alongside a representation of the conventional rawinsonde network, we only assimilate PECAN rawinsondes closest to 0000 UTC in all experiments. If no rawinsondes are launched from a site between 2330 and 0030 UTC, then all data from that site are excluded for that case. Finally, we assimilate all sets of collocated observations together in SONDE_ALLPROF.

4.3 System configuration and verification techniques

4.3.1 Model and DA system configuration

To simulate the 13 nocturnal CI events, we apply the multi-scale model and DA system described in Chapter 2.1-2.2. Although data gaps prevent the same quantity of data from being assimilated for each nocturnal CI case, we ensure that each site is at least assimilated for the same duration by always beginning DA 16.5 h prior to the primary, observed CI event (Fig. 4.1b; 13.5 h of DA and 3 h of forecast spin-up). Four assimilation cycles are first performed every 3 h on the outer, CONUS domain (Δx=12 km; d01 in Fig. 4.1) using conventional data, PECAN rawinsondes, and the PECAN remote sensing profilers to improve the analysis of synoptic and mesoscale features. Afterwards, we downscale the outer domain to an intermediate, convection-allowing grid where radar observations are also assimilated for six, 15-min cycles to improve analyses of storm-scale features (Δx=4 km; d02 in Fig. 4.1). Previous studies (e.g., Johnson and Wang 2019) indicate that at least 1-km grid spacing is necessary to capture small-scale convective events such as those initiating along an undular bore. Thus, after DA is complete, we initialize 7-h forecasts from the final EnKF analysis on a third, 1-km forecast domain (see new d03 in Fig. 4.1). Following Johnson and Wang (2019), we only use the first 10 ensemble members for the forecast period to represent current convection-allowing ensemble systems.

4.3.2 Object-based identification of CI

Given that CI is a locally rare and rapidly occurring process (Burlingame et al. 2017), applying traditional verification methods can be difficult (e.g., Kain et al. 2013). Recently, Burghardt et al. (2014) developed an object-based technique that allows for detecting and matching between specific observed and forecast CI objects. We apply this method, also utilized by Burlingame et al. (2017) and Keclik et al. (2017), to identify each CI event occurring during the 7-h forecast period (example shown in Fig. 4.3). The same technique is simultaneously applied to both the forecasts and MRMS radar observations (Smith et al. 2016) using the following steps. First, we apply bilinear interpolation to regrid the observed MRMS reflectivity data onto the model grid for each case. We then identify convectively active regions in the forecast and observations by selecting areas of reflectivity greater than 35 dBZ at the -10 °C level. Searching at the -10 °C isotherm avoids potential brightbanding effects that could cause spurious CI events to be detected in stratiform precipitation regions (Gremillion and Orville 1999). For the forecast data, we calculate reflectivity at -10 °C by searching downwards from the model top to find the first model level where the ambient temperature is greater than -10 °C. The forecast reflectivity data are then interpolated onto that height. The MRMS reflectivity at -10 °C is calculated using the same method, except that RAP analyses are used for determining the height of the -10 °C isotherm (NOAA 2015). We note that the height fields derived from the forecast data and RAP analyses are generally within 100 m of each other and have little impact on the derived reflectivity product.

Once convectively active regions have been identified, we detect individual convective objects by segregating the reflectivity data with a watershed transform algorithm that requires at least four contiguous grid points of reflectivity $>$ 35 dBZ (Lakshmanan et al. 2009). These

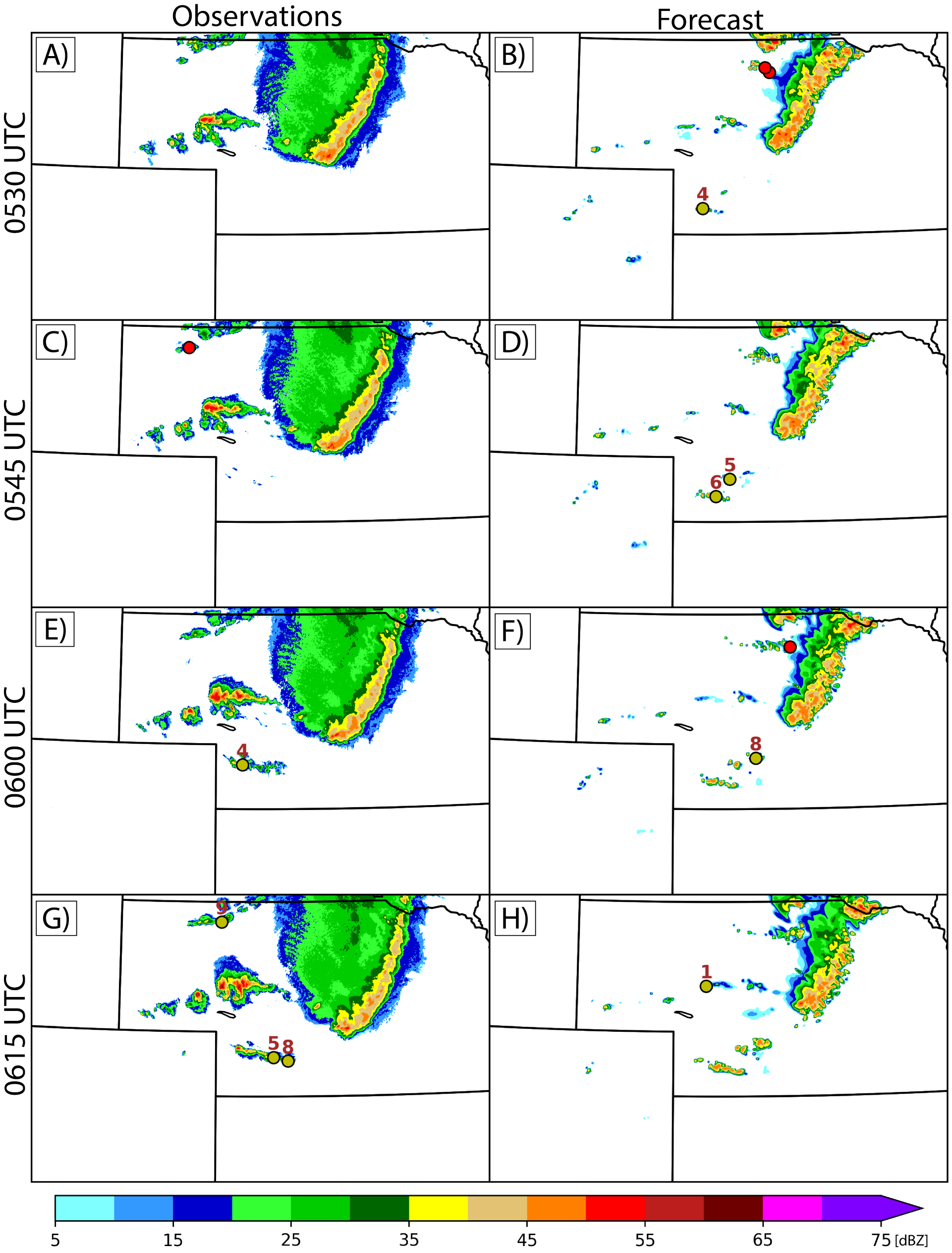


Figure 4.3: Example of the object-based CI detection algorithm applied throughout this study for forecast member 1 of SONDE_ALLPROF on 3 June. The background shading represents the (a,c,e,g) observed and (b,d,f,h) forecast reflectivity at -10 °C (dBZ). The unmatched CI objects, indicated at the time and centroid location of CI, are represented by red dots. If a CI object is matched between the forecast and observed objects, it is instead represented by a yellow dot and the pair ID is annotated above the object.

convectively active objects are then tracked backwards in time so that their starting time and location (CI) can be determined. The object tracking works by searching a 15-km search radius, which we subjectively find to work well for the cases simulated here, around each object at the previous output time (15 min). This step effectively tracks objects across time by merging any pairs with a translation speed less than 60 km h^{-1}. We note that only ∼2% of objects move faster than this speed, indicating that the threshold likely has little impact on the results presented herein. After tracking, we remove any objects that are not maintained for a minimum of 1 h, such that we only account for mature, deep convection (Burlingame et al. 2017). We then define the CI objects by the starting time and centroid location of these convectively active objects (red and yellow dots in Fig. 4.3).

4.3.3 Object-based verification of CI

To quantify the capability of each forecast experiment to predict CI, we apply both probabilistic and deterministic verification methods based on the CI objects detected using the previous method. The forecast probabilities are created using individual binary masks for each ensemble member. Given that CI often occurs at only a few grid points, we determine ensemble probabilities using unsmoothed neighborhood maximum ensemble probabilities (NMEP) that Schwartz and Sobash (2017) recommend for locally rare events such as updraft helicity tracks. NMEP measures the percentage of ensemble members that produce CI within a selected neighborhood. Once NMEP has been calculated for each case, we then verify the probabilistic forecast fields using FSS (Roberts and Lean 2008) which measures how well each experiment predicts general regions of nocturnal CI.

The deterministic verification technique (Fig. 4.3) matches between forecast and observed objects based on subjectively chosen spatiotemporal thresholds (Burghardt et al. 2014). We match between forecast and observed objects if (1) the centroid distance between the pair is less than 150 km, and (2) if the absolute difference in their initiation times is less than 2 h. We also perform various sensitivity tests that change these spatiotemporal thresholds and find little to no impact on the relative performance between the assimilation experiments. If we find multiple potential matches between object pairs, then we match the pair with the lowest composite error (E; Burghardt et al. 2014), defined as:

$$E = [E_d^2 + (V_c * E_t)^2]^{1/2} \tag{4.1}$$

where E_d is the centroid distance error, E_t is the absolute timing error, and V_c is a representative storm motion calculated by averaging the simulated wind speed between 450 and 850 hPa at the location of the observed CI object (Burlingame et al. 2017).

After matching between the forecast and observed objects (yellow dots in Fig. 4.3), we organize results into a 2 X 2 contingency table and calculate the standard verification metrics (Wilks 2011) of probability of detection (POD), false alarm ratio (FAR), bias, and critical success index (CSI). Recent studies (e.g., Skinner et al. 2016; Schwartz and Sobash 2017) explore various ways to compute similar object-based verification metrics across an ensemble. Here, we compute these contingency metrics by combining hits, misses, and false alarms from each forecast member into a single contingency table such that one value of each metric is computed for the ensemble per case. Consequently, the limited number of cases (n=13) prevents the estimation of statistical significance. In addition to the contingency metrics, we also calculate timing, distance, speed, direction, aspect ratio, and axis angle errors of the matched object pairs following the definitions in Davis et al. (2006). We calculate the latter four attributes 1 h after CI such that the convection has had some time to evolve from its original cellular shape.

4.4 Systematic impact of assimilating remote sensing profiles for forecasts of nocturnal CI

We first assess the general impacts of assimilating the PECAN profilers by verifying NMEP forecasts of both hourly precipitation exceeding 6.35 mm (0.25 in) and CI (Figs.

4.4, 4.5). In general, each experiment performs similarly for precipitation with only slight changes to the FSS when assimilating remote sensing data. Compared to SONDE, assimilating the thermodynamic profilers in SONDE_TQPROF slightly improves the mean precipitation skill at most forecast hours and for both neighborhood sizes (Fig. 4.4a,b). Most of the impact to the mean FSS skill in SONDE_TQPROF is a result of three cases (i.e., outliers in Fig. 4.5a,b). Conversely, assimilating the kinematic profilers leads to a consistent negative impact for precipitation. SONDE_UVPROF produces lower skill compared to SONDE, especially during the early and later forecast hours (Fig. 4.4a,b), leading to median FSS impacts of -0.02 to -0.01 depending on the exact neighborhood size (Fig. 4.5a,b). When assimilating both sets of profilers together in SONDE_ALLPROF, we find little impact to the FSS such that the negative impacts of assimilating the composite kinematic profilers likely counteract the positive impacts of assimilating the thermodynamic profilers. Additionally, though the forecasts are more skillful when verifying over a larger neighborhood, the impacts remain generally the same magnitude at both r = 50, 150 km.

Similar to results found by Keclik et al. (2017), all experiments produce much lower skill for CI forecasts compared to precipitation (Figs. 4.4,4.5). Though not necessarily surprising, this suggests that convection-allowing NWP systems can predict general regions where precipitation is likely to occur, but struggle to predict the precise timing and location of initiation within those regions. We also find that assimilating the PECAN profilers has a larger overall impact for CI compared to precipitation (Fig. 4.4c,d; Fig. 4.5c,d). Using a 50-km verification neighborhood, we find that SONDE_TQPROF produces increased skill compared to SONDE at ∼4.5 h (Fig. 4.4c), and between 3.5-6.5 h when using a 150-km neighborhood (Fig. 4.4d). Conversely, SONDE_UVPROF often produces similar or lower skill compared to SONDE, most notably at ∼3 h using a 50-km neighborhood (Fig. 4.4c) and at ∼5 h when using a 150-km neighborhood (Fig. 4.4d). Again, these impacts likely cancel each other out when both datasets are assimilated, such that SONDE_ALLPROF produces similar skill to SONDE (Fig. 4.4c,d 4.5c,d). We discuss hypotheses for why

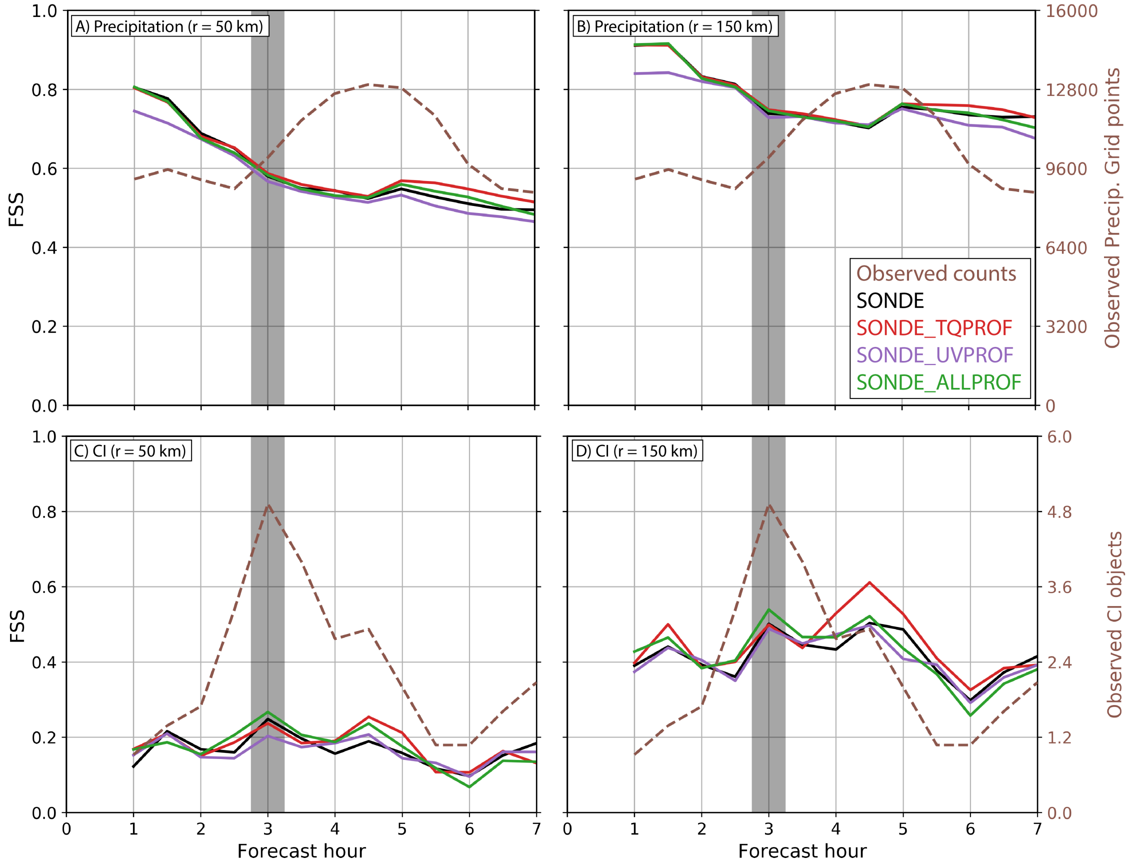


Figure 4.4: Time series of fractions skill score (FSS) for varying neighborhoods. The time series is created by averaging the time series of FSS for the 13 individual CI cases. The statistics are calculated for (a,b) hourly precipitation exceeding 6.35 mm (0.25 in); and (c,d) CI objects detected over the previous hour. Also shown in (a,b) are the average number of grid points where observed hourly precipitation exceeds 6.35 mm, and (c,d) the average number of observed CI objects. While each CI event during the 7 h forecast is verified here, the timing of the primary CI event used to select the 13 cases is indicated by the gray shading.

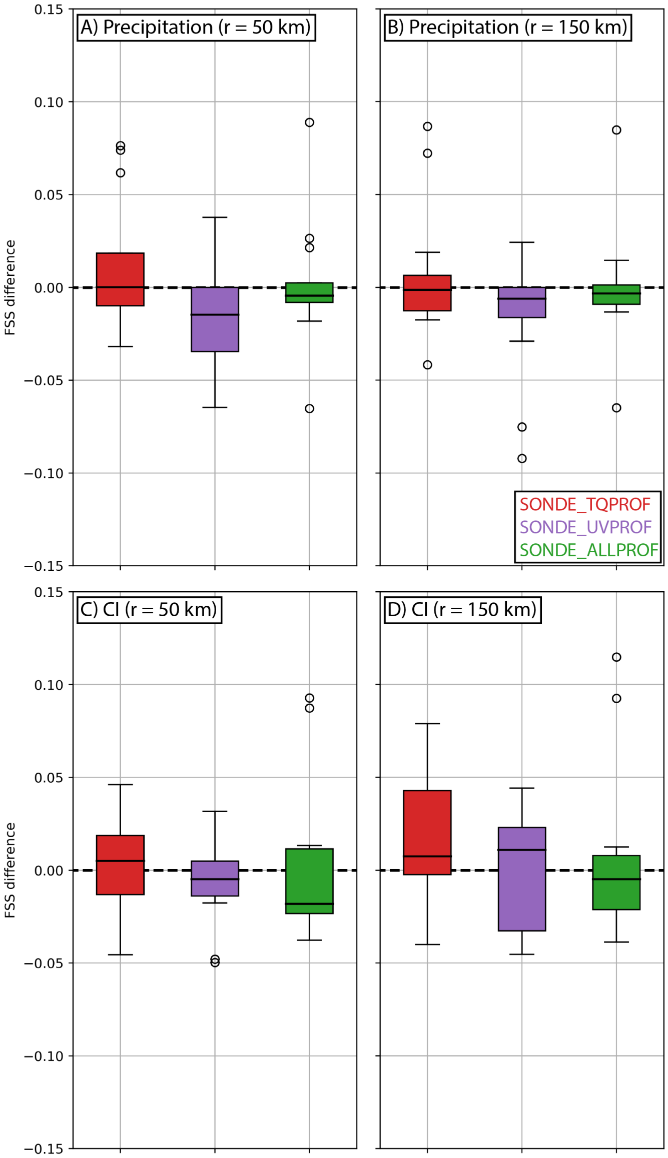


Figure 4.5: Box-and-whisker plots for FSS differences between the indicated experiment and SONDE ($\mathrm{FSS}_{\mathrm{EXPERIMENT}} - \mathrm{FSS}_{\mathrm{SONDE}}$) using varying neighborhoods. The statistics are calculated for (a,b) hourly precipitation exceeding 6.35 mm (0.25 in); and (c,d) CI objects detected over the previous hour. The line through the middle of the box represents the median of the data. The bottom of the box represents the lower quartile and the top of the box represents the upper quartile. The whiskers extend to the 1.5 times the interquartile range (IQR). Any outliers above or below 1.5*IQR are plotted as circles.

assimilating kinematic data sometimes degrades the nocturnal convective forecasts in later sections. Additionally, the FSS for CI shows stronger case-to-case variability (as indicated by the taller boxes in Fig. 4.5c,d) and more hourly variability than the FSS for precipitation (Fig. 4.4c,d). The latter result is likely related to the large variability in the number of observed CI objects, as Fig. 4.4c,d shows local minimums in skill when the observed object counts decrease at forecast hour 6. These increased impacts for CI compared to precipitation suggests that nocturnal CI forecasts are more sensitive to the boundary layer modifications generated by assimilating the remote sensing data. While small changes to the low-level pre-convective environment might slightly shift the location or magnitude of large-scale precipitation regions, similar changes can also cause CI forecasts to entirely fail or new CI events to occur (e.g., Martin and Xue 2006).

While the FSS metric quantifies how well each experiment predicts general regions of CI, we also perform a deterministic verification technique that matches between specific observed and forecast CI events (Table 4.4). This method allows us to not only compute the percentage of CI events that are successfully predicted, but also measure the number of false alarms and estimate various errors of the successful matches. We note that the "POD-primary" in Table 4.4 evaluates only over objects associated with the primary CI events used to select the 13 CI cases (Table 4.2), while the other detection statistics verify over every CI object within the model domain.

Similar to the probabilistic verification results, we find that each experiment produces similar contingency metrics with mean PODs and FARs of $\sim$0.50, biases of $\sim$1.10[1], and CSIs of $\sim$0.33 (Table 4.4). While the assimilation impacts to these contingency metrics are small, we find measurable impacts that agree with the findings presented for the probabilistic verification methods. First, assimilating the thermodynamic retrievals in SONDE_TQPROF produces the best results for many contingency metrics. Compared to SONDE, SONDE_TQPROF leads to small increases in POD (+0.01), POD-primary (+0.04),

---

[1] The mean bias for each experiment is greater than 1.0 despite a lower average number of forecast objects compared to observed objects. This is due to most cases featuring a bias less than 1.0, but three cases producing many forecast objects where the individual bias for that case is greater than 1.6.

Table 4.4: Mean contingency statistics for each experiment including the number of observed and forecast CI objects per case, probability of detection (POD), false alarm ratio (FAR), bias, and critical success index (CSI). The number of forecast objects is also averaged across the ten ensemble members. POD, FAR, bias, and CSI are calculated over every nocturnal CI event within the forecast domain, while POD-primary is calculated over only the nocturnal CI events corresponding to the primary ascent mechanisms listed in Table 4.2.

Statistics	SONDE	SONDE_TQPROF	SONDE_UVPROF	SONDE_ALLPROF
# observed CI objects	15.5	15.5	15.5	15.5
# forecast CI objects	15.4	15.3	14.2	15.3
POD	0.53	0.54	0.50	0.52
FAR	0.48	0.48	0.47	0.50
Bias	1.13	1.09	1.04	1.14
CSI	0.34	0.35	0.33	0.33
POD-primary	0.66	0.70	0.64	0.68

and CSI (+0.01), and a small decrease in the bias (-0.04). Conversely, assimilating the kinematic profilers produces the worst overall metrics for POD, CSI, and POD-primary. SONDE_ALLPROF performs similarly to SONDE, indicating that positive and negative impacts of assimilating the two datasets again counteract each other. We note the average POD is slightly higher when verifying over objects associated with the primary CI events targeted by the data assimilation (∼0.64). In addition to being located closer to the mobile observations, these events are often the largest within the domain and are likely easier to capture. However, the impact on the rest of contingency statistics generally do not change when limited to the primary CI events (not shown).

We also present performance diagrams (Roebber 2009) for each individual CI case in Fig. 4.6. In general, we find strong case-to-case variability in both the baseline performance and for the impact of assimilating the remote sensing profilers. For example, SONDE_ALLPROF shows a large increase in CSI compared to SONDE for the 24 June frontal overrunning case (pentagons in Fig. 4.6). However, assimilating both datasets together has the opposite impact for the 26 June frontal overrunning case (stars in Fig. 4.6). Likewise, we also find a few cases where assimilating the thermodynamic data degrades the forecast skill (e.g., 16 July given by triangles in Fig. 4.6) despite the mean contingency impacts being positive. Romine et al. (2016) find similar results when assimilating dropsonde data, noting that the cases with the most positive impacts feature observations which sample broad regions around the event of interest. Thus, the case-to-case variability shown here is likely sensitive to both the number of observing sites assimilated and their location relative to the CI event (see marker size and transparency in Fig. 4.6). We also compare the impact of assimilating the remote sensing data to the strength of synoptic-scale ascent (Fig. 4.7). Cases that feature an upstream trough at either 500 or 700 hPa are classified as strongly-forced (6 out of 13 cases), while other events are classified as weakly-forced (7 out of 13 cases). This subjective analysis reveals that most of the impact to the mean CSI in SONDE_TQPROF results from assimilating the AERI data for strongly-forced cases (Fig. 4.7b). Instead, assimilating the AERI observations for weakly-forced cases produces little-to-no impact, suggesting that much of the case-to-case variability could also be related to the synoptic regime that the nocturnal CI events develop within.

Finally, we also calculate mean attribute errors for the successful matched pairs (Table 4.5). We note that the timing and distance errors shown are likely biased, as an event would only be included in this calculation if it passed certain thresholds of these errors to be considered a match. Overall, each experiment produces near zero mean timing errors. Although this result agrees with previous evaluations of CI forecasts (Johnson and Wang 2017; Kain et al. 2013), we note that the mean absolute timing error for each experiment is ∼45 min (not shown), indicating that NWP forecasts still struggle to predict the precise timing of CI. Each experiment also produces a distance error of ∼70 km which is slightly larger than previous studies (e.g., ∼50 km in Keclik et al. 2017), though these exact values are likely sensitive to the specific ascent mechanism for each CI event. When assimilating the thermodynamic profilers, the experiments show small improvements to the distance, direction, and axis angle errors. Assimilating the composite kinematic profilers also slightly improves the timing, distance, aspect ratio, and axis angle errors. When both datasets are assimilated in SONDE_ALLPROF, the forecasts produce the lowest errors in timing, distance, direction, and axis angle. Thus, while we find that the impacts of assimilating the

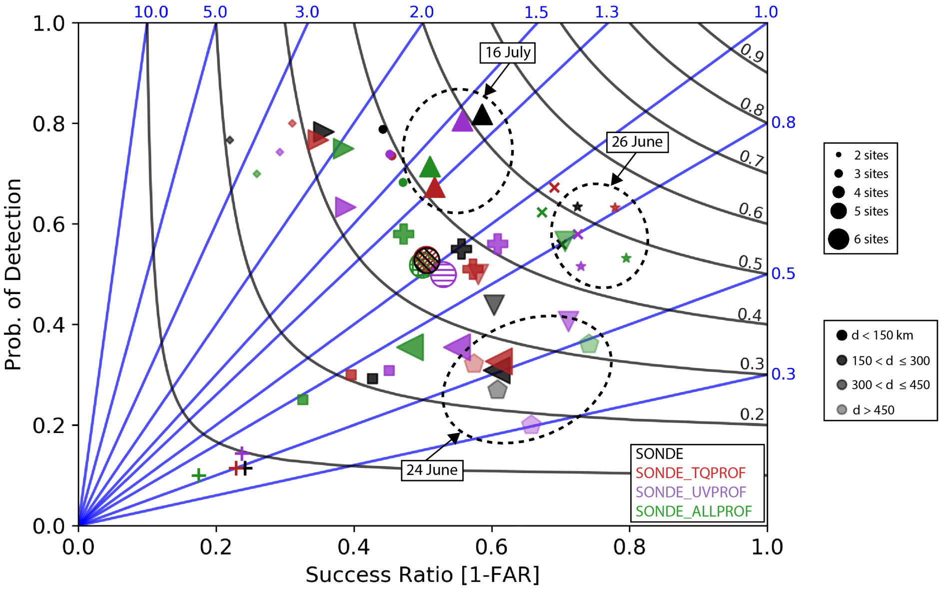


Figure 4.6: Performance diagrams for each nocturnal CI case. The x-axis is the success ratio (1 – false alarm ratio) and the y-axis is the probability of detection. The critical success index is represented by the curved black lines and the bias is represented by the straight blue lines. The different marker styles indicate a different nocturnal CI event, while the marker size indicates the number of sites assimilated and the transparency of the maker indicates the average distance between the geographic center of the observing sites and the primary CI event of interest (Table 4.2). Additionally, the hatched, circular markers near the center of the plot represent the mean across all 13 nocturnal CI cases.

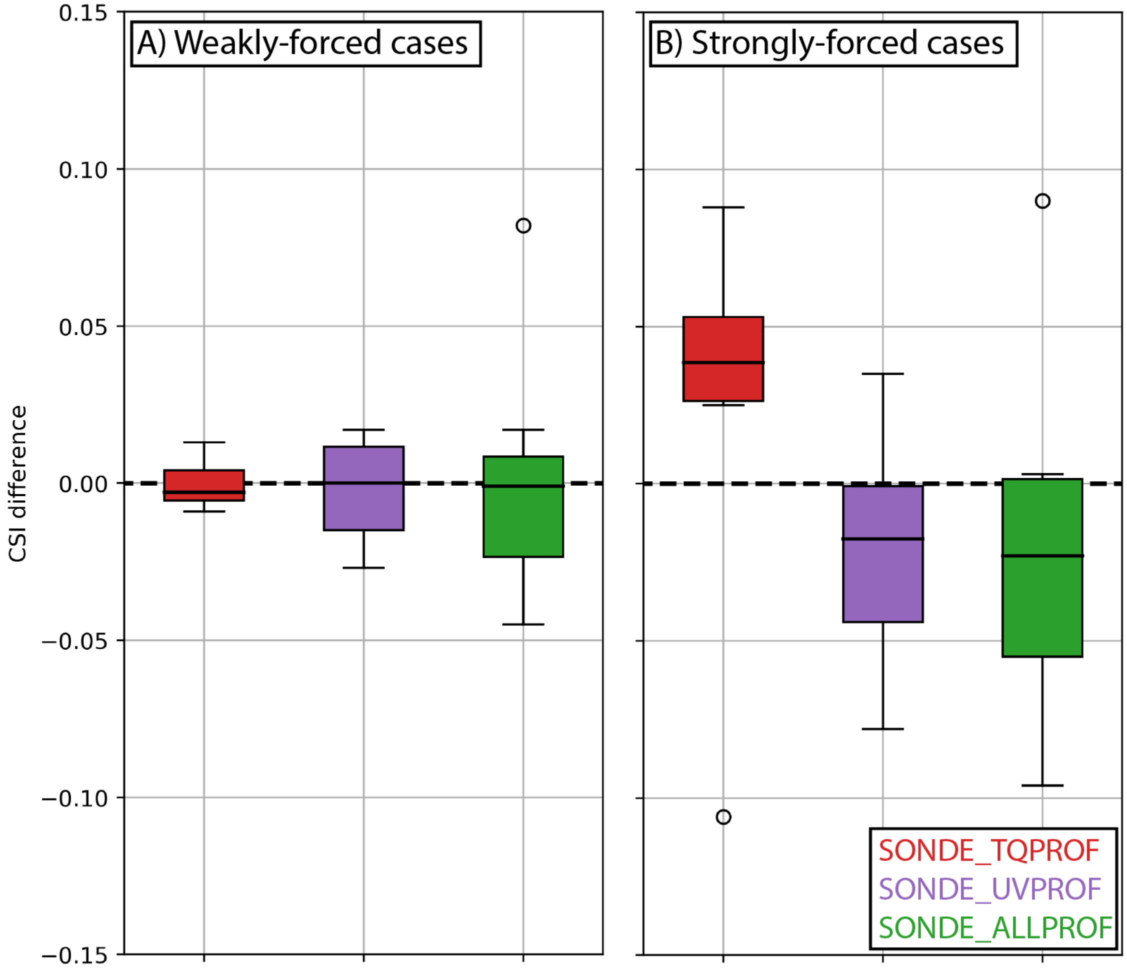


Figure 4.7: As in Fig. 4.5 but for CSI differences between the indicated experiment and SONDE ($\mathrm{CSI}_{\mathrm{EXPERIMENT}} - \mathrm{CSI}_{\mathrm{SONDE}}$) for (a) weakly-forced; and (b) strongly-forced cases.

Table 4.5: Mean attribute errors for the successful matches (hits) of each nocturnal CI event. All errors are calculated as (*forecast attribute – observed attribute*) and all errors except for distance can be either positive or negative. The timing and distance errors are calculated at the valid time for CI, while the speed, direction, aspect ratio, and axis angle errors are calculated 1 h after CI such that the convection has had some time to develop from its original cellular shape.

Mean error	SONDE	SONDE_TQPROF	SONDE_UVPROF	SONDE_ALLPROF
Timing (min)	2.5	3.5	-1.1	0.2
Distance (km)	70.4	68.9	69.2	65.5
Speed (ms^{-1})	-0.2	-0.2	-0.4	-0.1
Direction (°)	13.4	12.4	14.3	12.4
Aspect ratio	0.04	0.04	0.02	0.07
Axis angle (°)	-9.6	-9.0	-8.3	-2.6

two profiler networks counteract each other for detection of CI, we instead find improvements to the attributes of the CI event when assimilating both thermodynamic and kinematic data together.

4.5 Pre-convective analysis of observation impacts

To better understand why assimilating the thermodynamic profilers results in better forecasts of nocturnal CI and precipitation, and why assimilating the kinematic datasets often degrades these forecasts, we create composite time-height cross sections leading up to the CI event. The cross sections are formed by averaging each case and ensemble member over a 75-km radius centered on the primary observed CI event to capture the average distance error between forecast and observed objects (Table 4.5). We then plot differences in the convective ingredients (Johns and Doswell 1992) between SONDE and other experiments in Figs. 4.8-4.10.

In general, separately assimilating the collocated profilers has similar impacts on the midlevel divergence (Fig. 4.8a,b). Both the thermodynamic and kinematic profilers enhance the midlevel ascent during the 2 h period prior to CI (i.e., reduced divergence in Fig. 4.8a,b), though SONDE_TQPROF enhances the ascent in a slightly higher layer (1.5-3 km AGL) compared SONDE_UVPROF (1-2 km AGL). We note that the increased low-level divergence in all figures $\sim$1 h prior to CI is likely a result of stronger cold pools and their associated downdrafts compared to SONDE. As Wilson and Roberts (2006) show that most nocturnal convective events develop in environments featuring midlevel convergence, the impacts shown here indicate that assimilating both sets of profilers generally yields a pre-convective environment that is more supportive of nocturnal CI. When both sets of profilers are assimilated together, SONDE_ALLPROF (Fig. 4.8c) shows an elevated divergence field similar to the individual experiments. Therefore, both datasets likely improve similar deficiencies in the ascent forecast by SONDE such that there is no significant change when assimilating the profilers together.

While the remote sensing profilers make similar modifications to the ascent prior to CI, we instead find large differences in their impact on the pre-convective thermodynamic fields (Fig. 4.9). For example, assimilating the AERIs results in a moister cross section above 500 m AGL (+0.3 $g\ kg^{-1}$; Fig. 4.9a). Assimilating the composite kinematic profilers instead produces drying throughout the same layers (-0.3 $g\ kg^{-1}$; Fig. 4.9b). We also find that assimilating the AERIs results in weak low-level warming and midlevel cooling (Fig. 4.9a), whereas assimilating the composite kinematic profilers leads to slight midlevel warming throughout most of the cross section (Fig. 4.9b). The connection between the thermodynamic modifications and the detection of nocturnal CI is clear when comparing the assimilation impacts for convective indices (Fig. 4.10). The cooling and moistening in SONDE_TQPROF corresponds to increased CAPE (Fig. 4.10a) by $\sim$100 $J\ kg^{-1}$ for elevated parcels originating between 500-1000 m AGL, and slightly reduced inhibition (Fig. 4.10a) by $\sim$5 $J\ kg^{-1}$ throughout the same layer. These modifications also correspond to $\sim$150 m of less lifting required for parcels initiating between the surface and 1 km AGL to reach their level of free convection (Δz_{LFC}; Fig. 4.8a), though Δz_{LFC} is most strongly modified 1-3 h prior to the observed CI time. As such, the modifications made in SONDE_TQPROF result in more buoyant parcels that require less lifting to reach their LFC and are thus more likely to produce CI compared to the baseline SONDE experiment. Conversely, SONDE_UVPROF

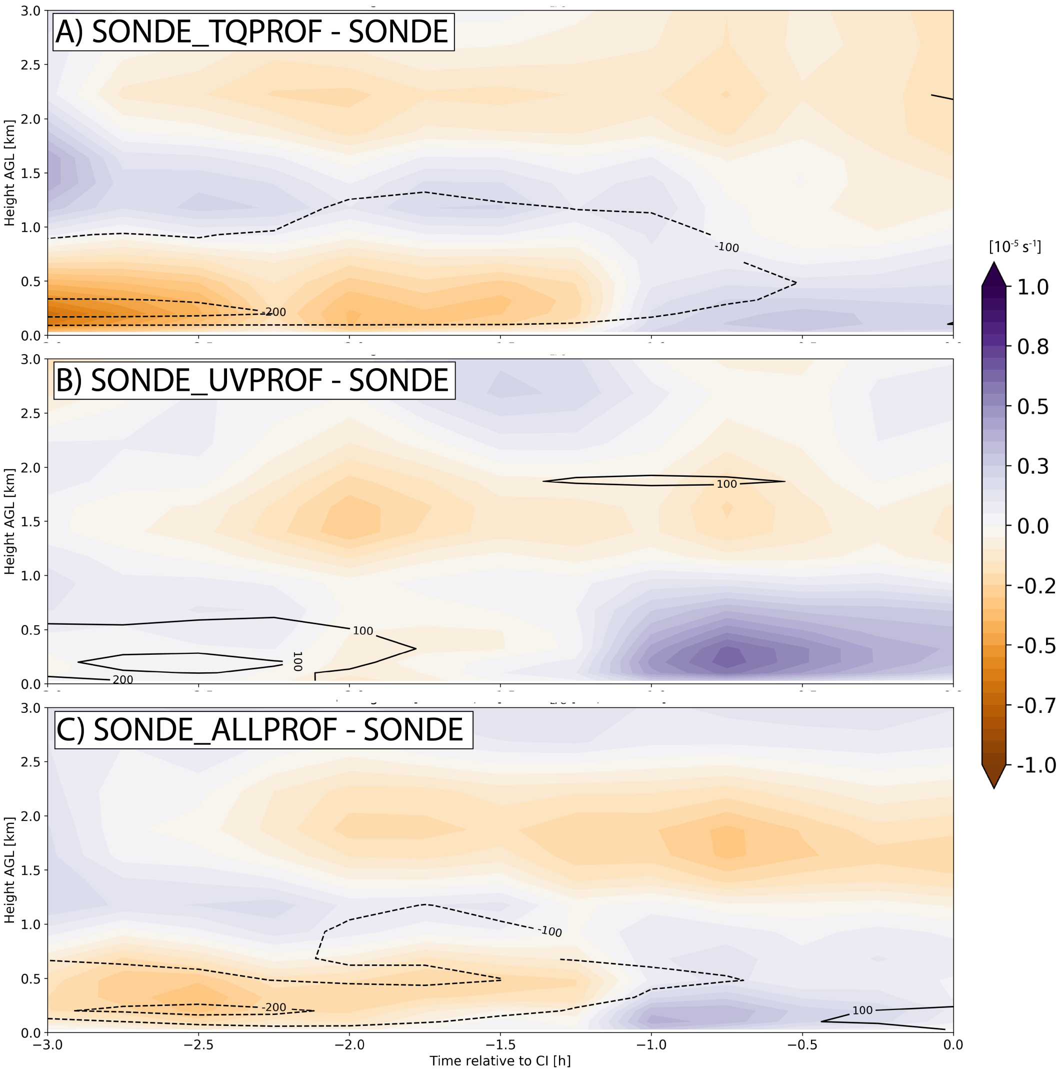


Figure 4.8: Composite, or mean, time-height cross sections for differences in divergence ($10^{-5}\ s^{-1}$; shading) and Δz_{LFC} (m; contours) between SONDE and other experiments. The cross sections are averaged over a 75-km radius around the observed CI event. The cross sections are also plotted relative to the observed timing of CI such that the final point on the x-axis indicates CI.

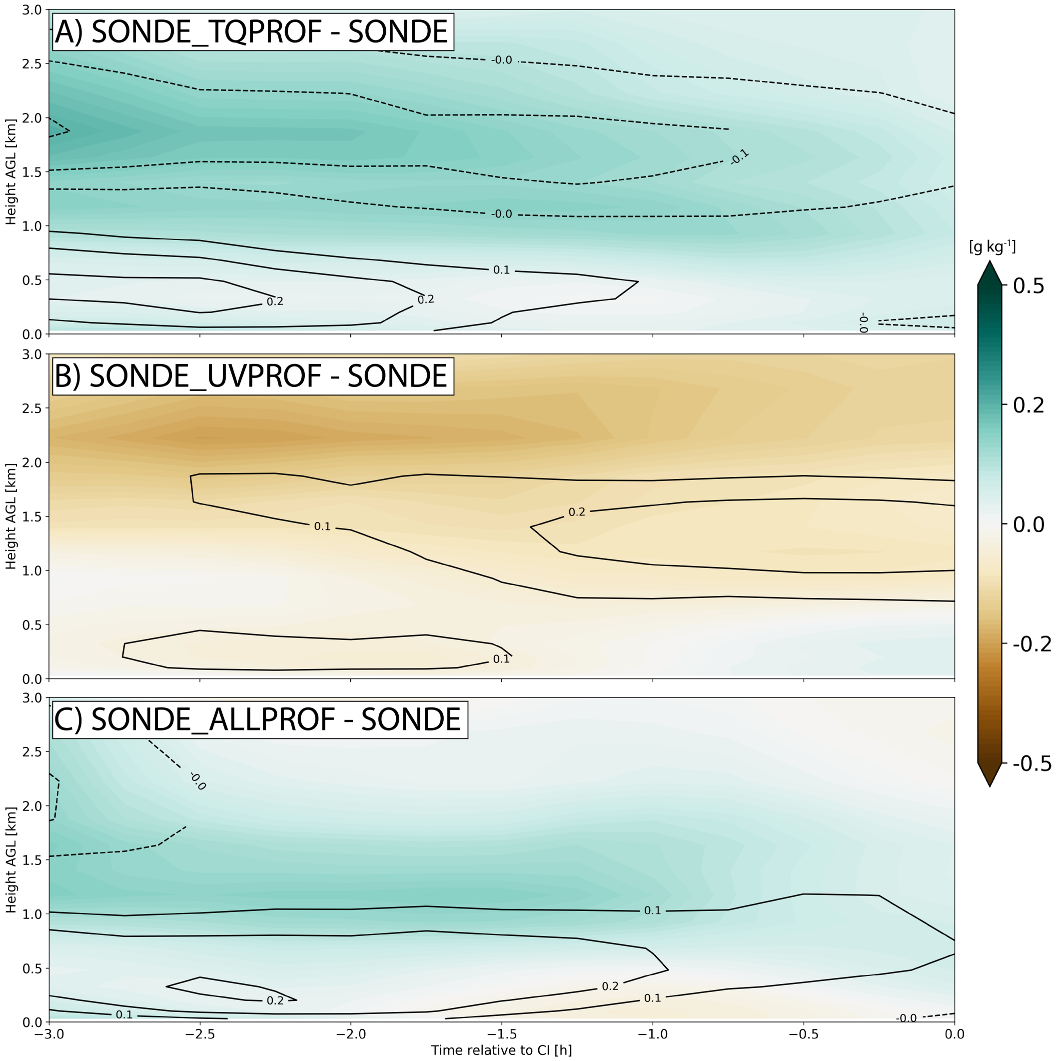


Figure 4.9: As in Fig. 4.8 but for differences in water vapor mixing ratio ($g\ kg^{-1}$; shading) and temperature (°C; contours).

results in modifications that are largely neutral including only a small increase in CAPE (Fig. 4.10b), almost no decrease in CIN (Fig. 4.10), and no reduction in Δz_{LFC} (Fig. 4.8b). These findings likely explain why assimilating the thermodynamic data results in more members being matched to the observed CI objects and overall better performance metrics, while assimilating the kinematic data likely degrades the nocturnal convective forecasts through updates to the pre-convective thermodynamic state that may hinder CI.

We perform various analyses to determine why assimilating the composite kinematic profilers often results in drying and warming prior to CI, while the AERIs, which directly

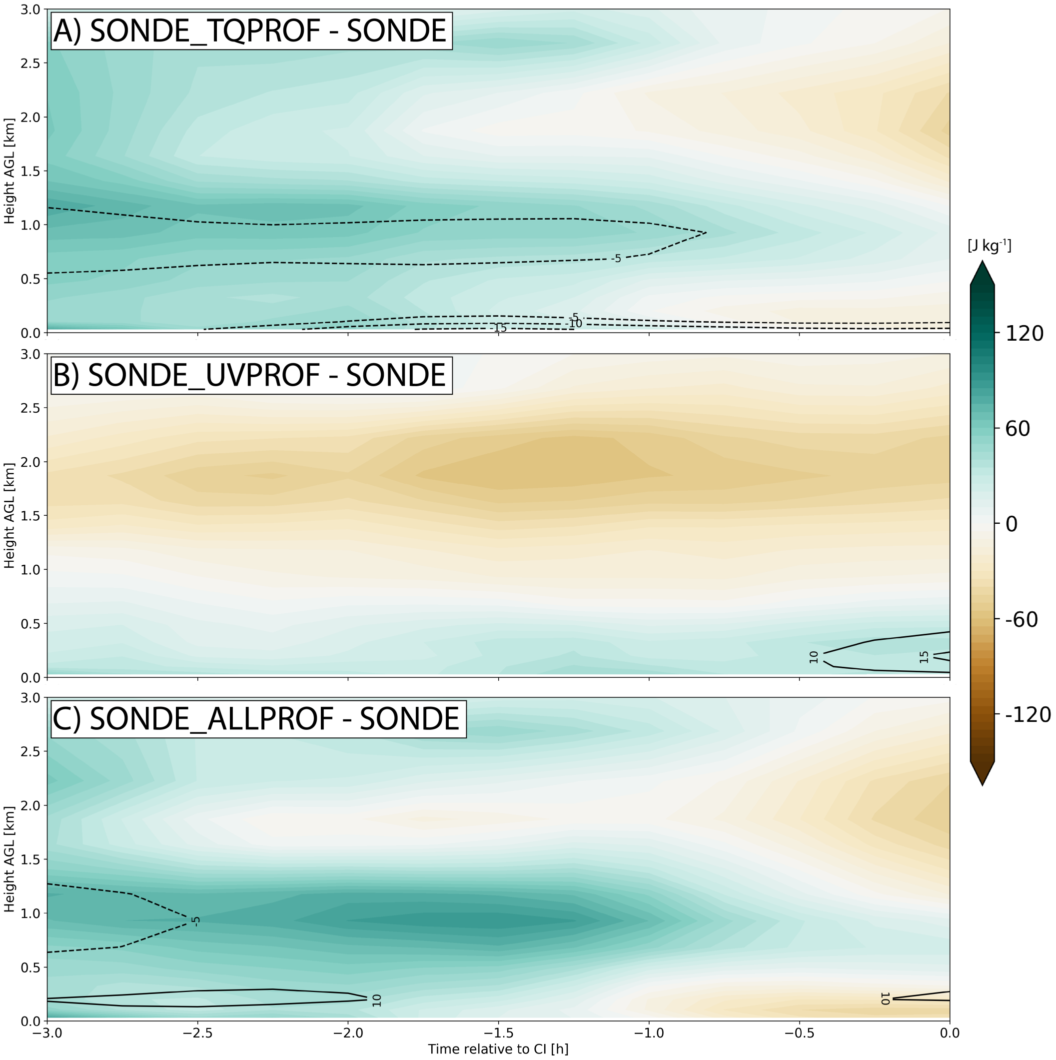


Figure 4.10: As in Fig. 4.8 but for differences in CAPE ($J\ kg^{-1}$; shading) and CIN ($J\ kg^{-1}$; contours). Negative values of CIN represent a decrease in inhibition for that experiment compared to SONDE.

measure these thermodynamic variables, produce opposite effects. In Fig. 4.11, we compare the assigned observation error standard deviations for the kinematic profilers at FP3 and FP6 to the RMSD between the profilers and collocated rawinsondes. Because the RSMD provides a measure of the instrument error, it should always be less than the assigned error (σ_c) that includes representation factors (Janjić et al. 2018). Although three of the six kinematic profiling sites feature reasonable observation errors (FP6 shown in Fig. 4.11a), three other sites show an RMSD that is often greater than the assigned observation errors (FP3 shown in Fig. 4.11b). The underestimated observation errors at FP3 become increasingly problematic above 1.5 km AGL and could lead to erroneous observations near the top of the wind profile being weighted more heavily than they should in the DA update. As such, it is possible that these underestimated kinematic errors at some sites could contribute to updates to the thermodynamic state that oppose the collocated AERIs. However, given that the representation component of observation errors is case- and site-dependent, simply increasing the errors by a constant amount might not improve forecast results for all cases. Advanced methods are therefore needed that can diagnose flow-dependent observation errors for kinematic profilers and remote sensing instruments in general.

4.6 Example cases

While assimilating kinematic profilers sometimes degrades the CI forecasts, three cases also show complementary results when assimilating both the kinematic and thermodynamic data together (Fig. 4.6). This final section explores two example cases in further detail, one which featured both datasets complementing each other such that SONDE_ALLPROF produces the best forecast of the nocturnal CI event, and one in which assimilating the composite kinematic data degrades the forecast.

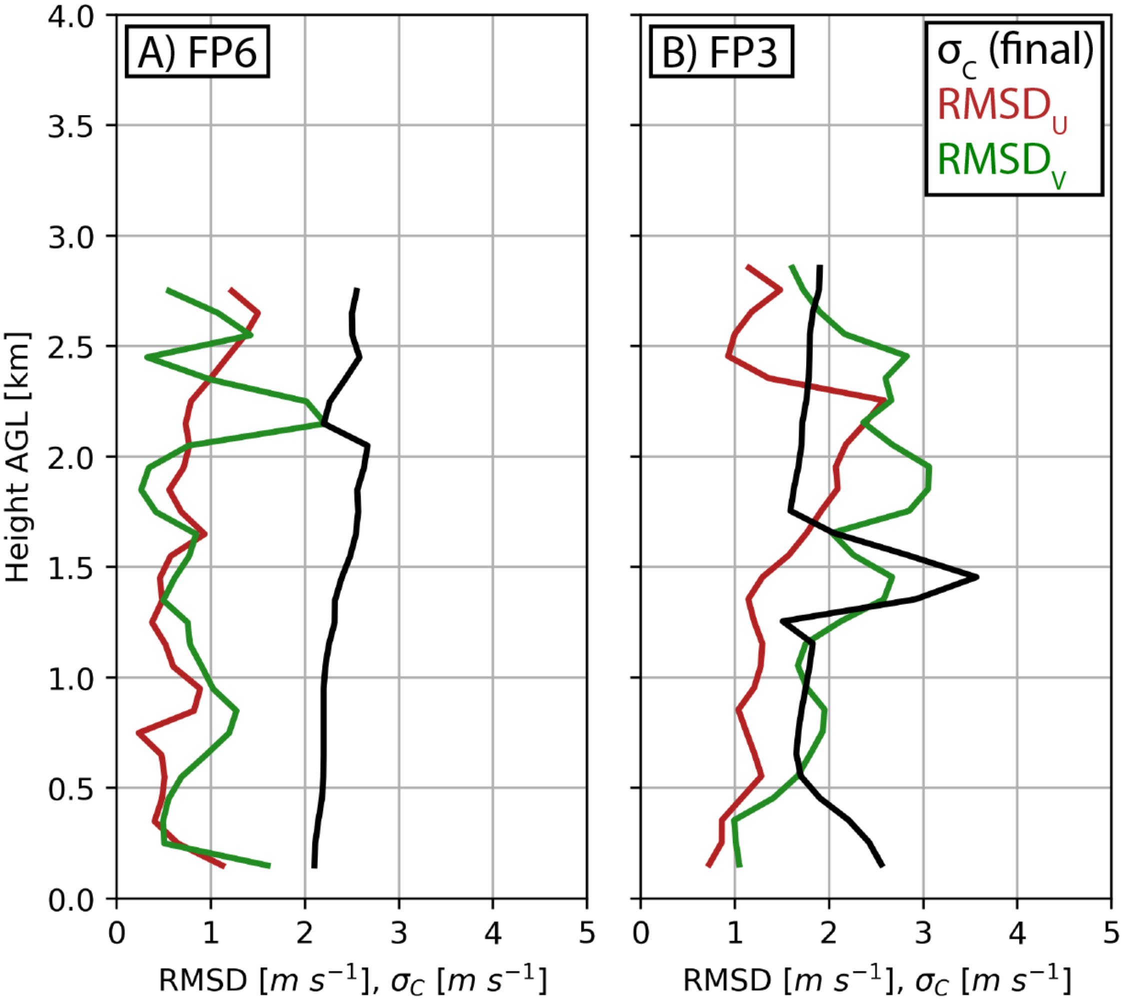


Figure 4.11: Root-mean-square differences (red and green; ms^{-1}) between the composite kinematic profilers and collocated soundings launched at (a) FP6 [n=10]; and (b) FP3 [n=21]. The location for each of these sites are shown in Fig. 4.1a. Also overlaid is the mean, inflated observation error standard deviation (black; ms^{-1}) for the same composite kinematic observations.

4.6.1 24 June frontal overrunning event

The 24 June event featured a large-scale CI episode that developed in central Iowa (dashed black oval in Fig. 4.12) near a region of isentropic ascent of the LLJ above a stationary surface boundary (Trier et al. 2020). The simulated CI was also partially generated by additional convergence both along the northern terminus of the LLJ, and along an outflow boundary generated by earlier convection in central Nebraska. Without any remote sensing data assimilated, SONDE predicts this event with low neighborhood ensemble probabilities (NEP; Schwartz and Sobash 2017) that are displaced too far south compared to the observed event at 0700 UTC (Fig. 4.12b). Assimilating the thermodynamic profilers increases the timing error for the initial CI event (Fig. 4.12a,d), but also correctly enhances the NEP values by $\sim$20% after the simulated CI occurs (Fig. 4.12c,f). We note that SONDE_TQPROF also enhances the NEP values of the spurious convection in northeastern Nebraska at 0930 UTC (dashed red oval in Fig. 4.12c,f). Assimilating the kinematic profiles reduces the overall extent of the predicted event and suppresses much of the spurious convection produced by SONDE and SONDE_TQPROF (Fig. 4.12i). When both remote sensing datasets are assimilated together, SONDE_ALLPROF predicts a convective system with large NEP values ($>$75%) that better match the observed event compared to other experiments (Fig. 4.12c,f,i,l). We note that the 24 June case features the largest positive impact from assimilating the remote sensing data despite the observing sites being over 500 km from the CI event on average (Table 4.2).

We find that assimilating the thermodynamic profilers primarily enhances the convective probabilities through increased midlevel moisture near the region of CI (Fig. 4.13). This moisture originates from AERI observations collected between 700 and 750 hPa at FP3 and FP5. These observations show large, positive innovations and lead to differences of +1-2 $g\ kg^{-1}$ in SONDE_TQPROF during the early assimilation cycles (Fig. 4.13a). The southwesterly background flow then advects this additional moisture into eastern Nebraska (Fig. 4.13c,e) and later into the region of CI. As a result, elevated parcels in

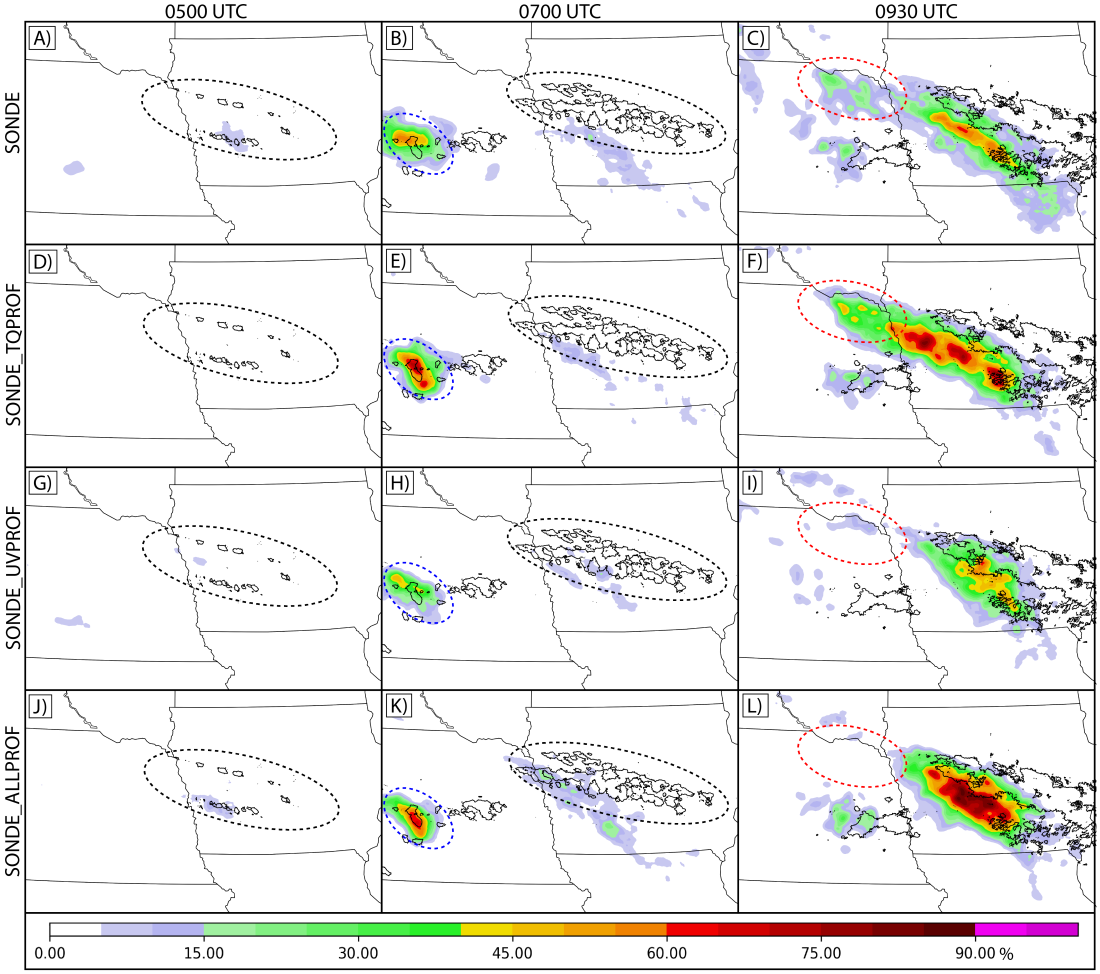


Figure 4.12: Neighborhood ensemble probability of composite reflectivity exceeding 30 dBZ for the 24 June frontal overrunning case valid at (a,d,g,j) 0500 UTC; and (b,e,h,k) 0700 UTC, and (c,f,i,l) 0930 UTC. Each plot is computed using an 8-km neighborhood. Also overlaid are the 30 dBZ contours of observed composite reflectivity. See text for a description of the dashed ovals.

SONDE_TQPROF feature increased CAPE and reduced CIN compared to SONDE (not shown). It is also likely that this additional moisture supports the enhanced probabilities for convection in central Nebraska (dashed blue ovals in Fig. 4.12e,k) that later generates the strong outflow boundary responsible for spurious CI events in northeastern Nebraska (dashed oval region in Fig. 4.15b).

Although assimilating the kinematic profilers results in slight drying near the region of CI (Fig. 4.13), we instead find that these data improve the CI forecast through an enhancement of the ambient wind speed in northeastern Nebraska (Figs. 4.14, 4.15). During early DA cycles, the FP3 and MP3 wind profilers observe faster wind speeds than the ensemble backgrounds at 800 hPa (shaded dots in Fig. 4.14b,d). Assimilating these data causes a large increase in the wind speeds around the Kansas and Nebraska border that is then spread throughout the region of CI (Fig. 4.14d,f). As such, SONDE_UVPROF and SONDE_ALLPROF produce stronger wind speeds in northeastern Nebraska by $\sim$2.5 ms^{-1} during the forecast period (Fig. 4.15c,d). When an outflow boundary produced by earlier convection in central Nebraska passes through this region at later forecast times, the enhanced background flow weakens the speed convergence (dashed oval region in Fig. 4.15) and thus the extent of the spurious convection (Fig. 4.12c,i,l). When both remote sensing datasets are assimilated together, their impacts complement each other such that SONDE_ALLPROF features the benefits of both the enhanced moisture from assimilating the thermodynamic data and the weakened convergence that reduces the spurious convection when assimilating kinematic data (Fig. 4.15d). These combined effects likely explain why SONDE_ALLPROF produces the most accurate convective forecast overall (Fig. 4.12l).

4.6.2 26 June frontal overrunning event

The nocturnal CI event on 26 June (also detailed in section 3.2) featured similar ascent mechanisms to the 24 June event, including frontal overrunning by the LLJ over a stationary boundary and enhanced convergence at the terminus of the LLJ. When assimilating the

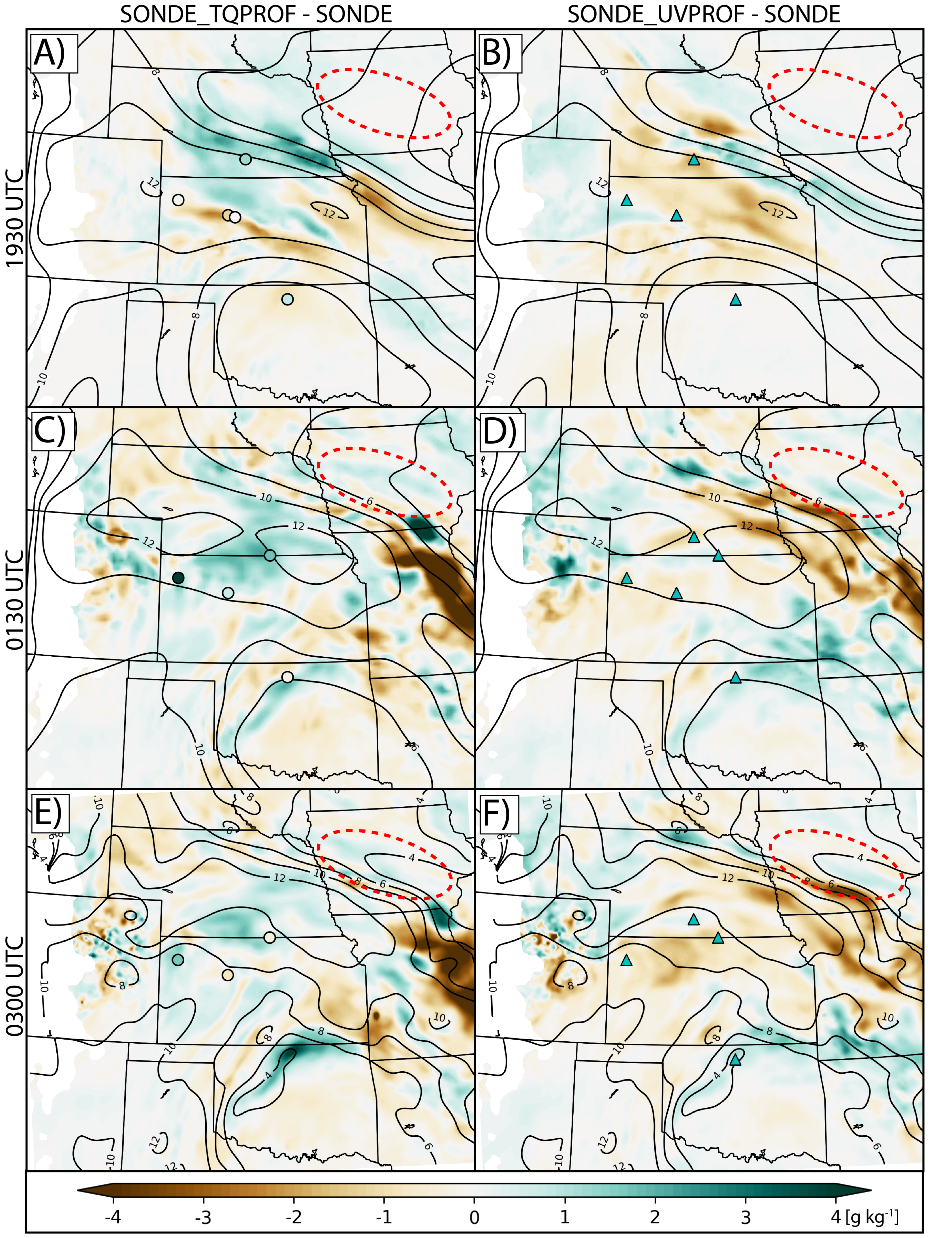


Figure 4.13: Differences in 700 hPa analyzed ensemble mean water vapor mixing ratio ($g\ kg^{-1}$) for (a,c,e) SONDE_TQPROF and SONDE; and (b,d,f) SONDE_UVPROF and SONDE. The plots are valid at (a,b) 1930 UTC 23 June; (c,d) 0130 UTC 24 June; and (e,f) 0300 UTC 24 June. Also contoured is the ensemble mean water vapor mixing ratio (every 2 $g\ kg^{-1}$) for SONDE. The shaded dots in (a,c,e) represent innovation values (observation minus background; same color and scale as fill) for the AERI observation closest to 700 hPa and assimilated in SONDE_TQPROF. The cyan triangles in (b,d,f) represent the location of composite kinematic profilers assimilated in SONDE_UVPROF. The dashed red ovals correspond to the primary location of the frontal overrunning CI event for 24 June.

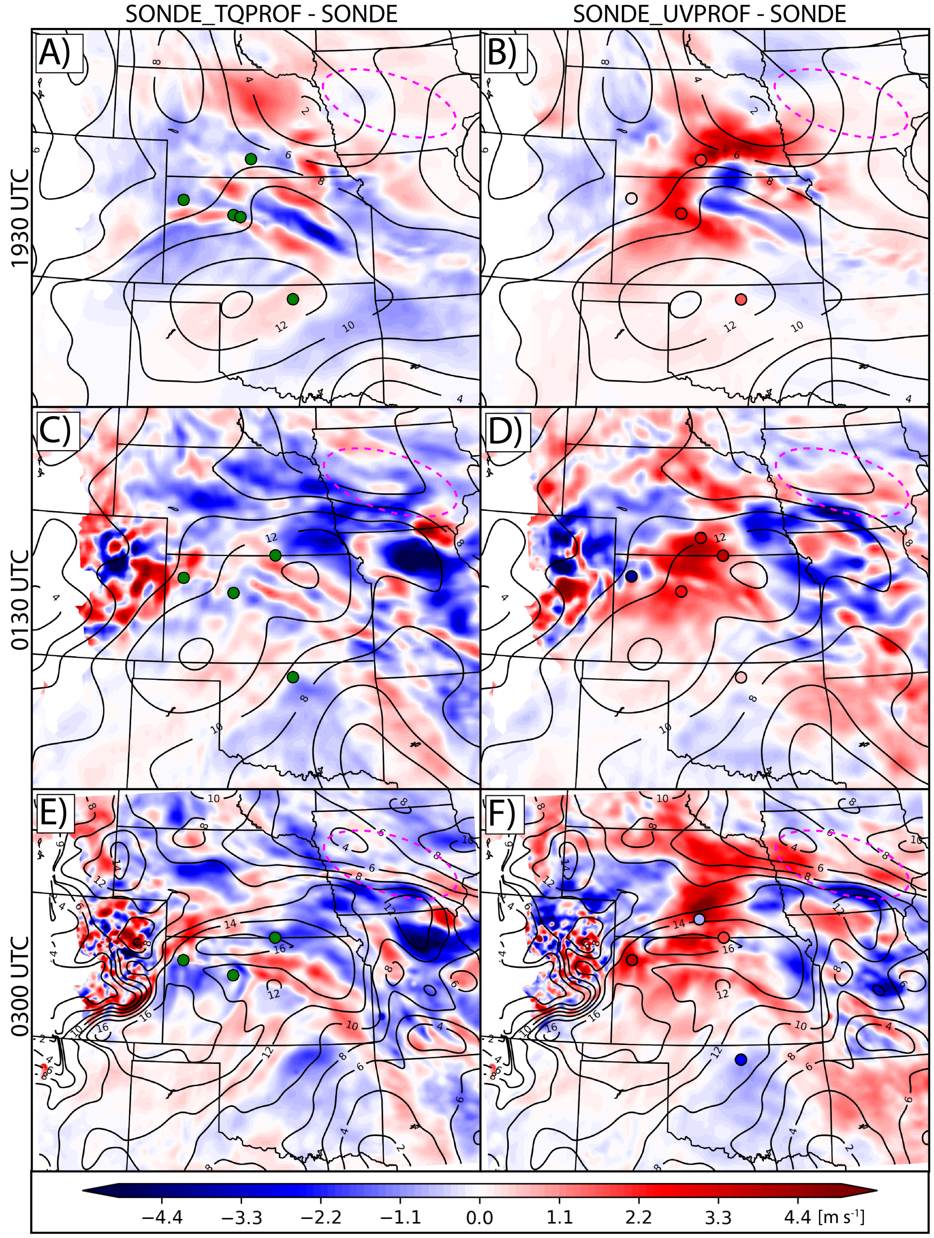


Figure 4.14: As in Fig. 4.13 but for differences in 800 hPa analyzed wind speed (ms^{-1}). Also contoured is the ensemble mean wind speed (every 2 ms^{-1}) for SONDE. The green dots in (a,c,e) represent the location of AERI observations assimilated in SONDE_TQPROF. The shaded dots in (b,d,f) represent the innovation values (observation minus background; same color and scale as fill) for the composite kinematic profiler observations closest to 800 hPa and assimilated in SONDE_UVPROF. The dashed pink ovals correspond to the primary location of the frontal overrunning CI event for 24 June.

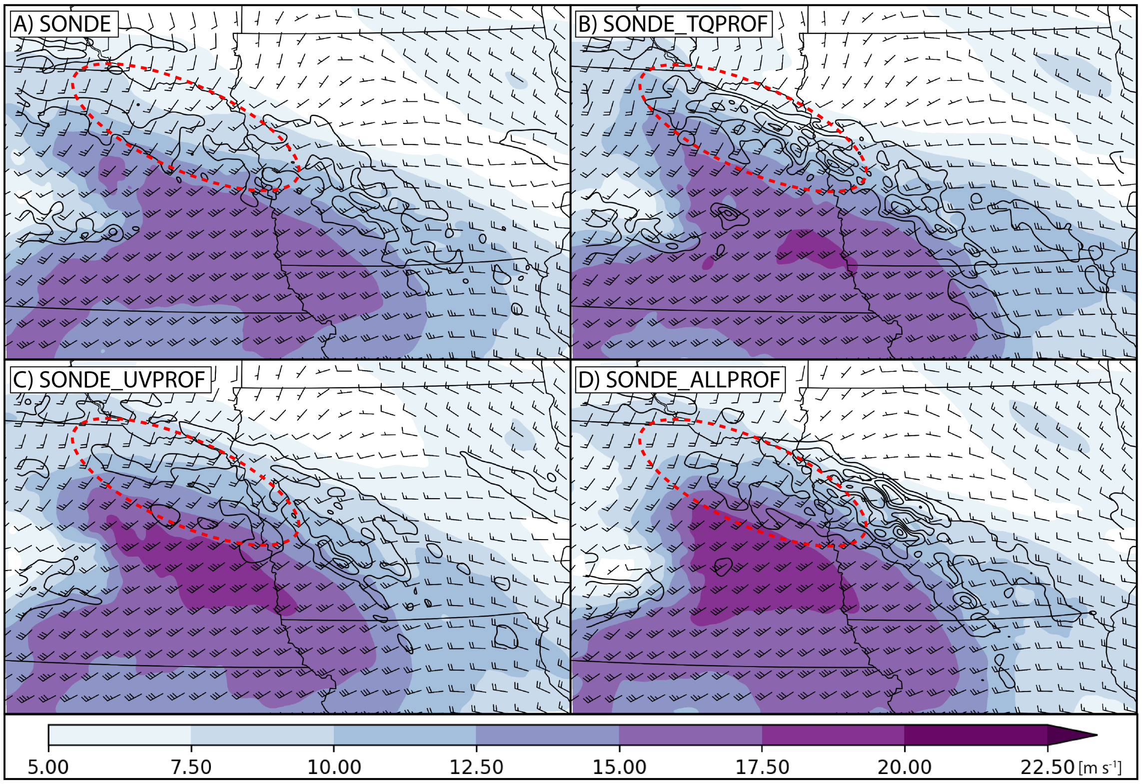


Figure 4.15: Ensemble mean forecasts of 800 hPa winds (ms^{-1}; shading and barbs) and horizontal convergence (contoured in black every +5^{-6} s^{-1}) valid at 0700 UTC 24 June. The half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. See text for a description of the dashed red ovals.

thermodynamic data, we find general improvements to the CI forecast compared to SONDE, such that higher NEPs are predicted along the central portion of the linear CI event and spurious convection to the southwest is suppressed (dashed black ovals in Fig. 4.16a,c). Conversely, SONDE_UVPROF predicts lower NEP values along most of the event (Fig. 4.16e). Additionally, at later forecast times when additional convection develops northwest of the main linear band, SONDE_UVPROF generates fewer convective cells and with lower probabilities compared to SONDE (dashed pink ovals in Fig. 4.16b,f). SONDE_ALLPROF also shows similar problems compared to SONDE_UVPROF such that the linear nocturnal CI event is predicted with a southeastern bias (Fig. 4.16g), and the second CI event is almost entirely missed (Fig. 4.16h).

As was discussed in the previous section, assimilating thermodynamic or kinematic data often has opposite effects on the pre-convective moisture environment. Fig. 4.17 demonstrates an example of this for the 26 June CI event. Assimilating the thermodynamic data introduces additional midlevel moisture (Fig. 4.17a), primarily originating from FP3, that is later advected into the region of CI and enhances the environment for new convective events (red dashed oval region in Fig. 4.17c,e). Conversely, assimilating the wind profiler data from the same site introduces drying that suppresses new development (Fig. 4.17b,d,f). We hypothesize that this drying in SONDE_UVPROF results from assimilating wind data near the top of the composite kinematic profile at FP3 (Fig. 4.18b). The wind observation at 1930 UTC is much slower than the background wind between 700-800 hPa and causes a large reduction in the analyzed v-wind when assimilated (Fig. 4.18b). Given the negative north-south moisture gradient near FP3 (Fig. 4.17b), assimilating this slower v-wind also weakens the northward moisture advection into central Kansas. Though we have no independent observations to verify the weakened moisture advection, the convective forecast serves as an indirect verification tool. Thus, compared to the moistening and improved CI forecast that results from assimilating the thermodynamic data, we assume that the data near the top of the wind profile at FP3 are likely incorrect. Without an accurate representation of

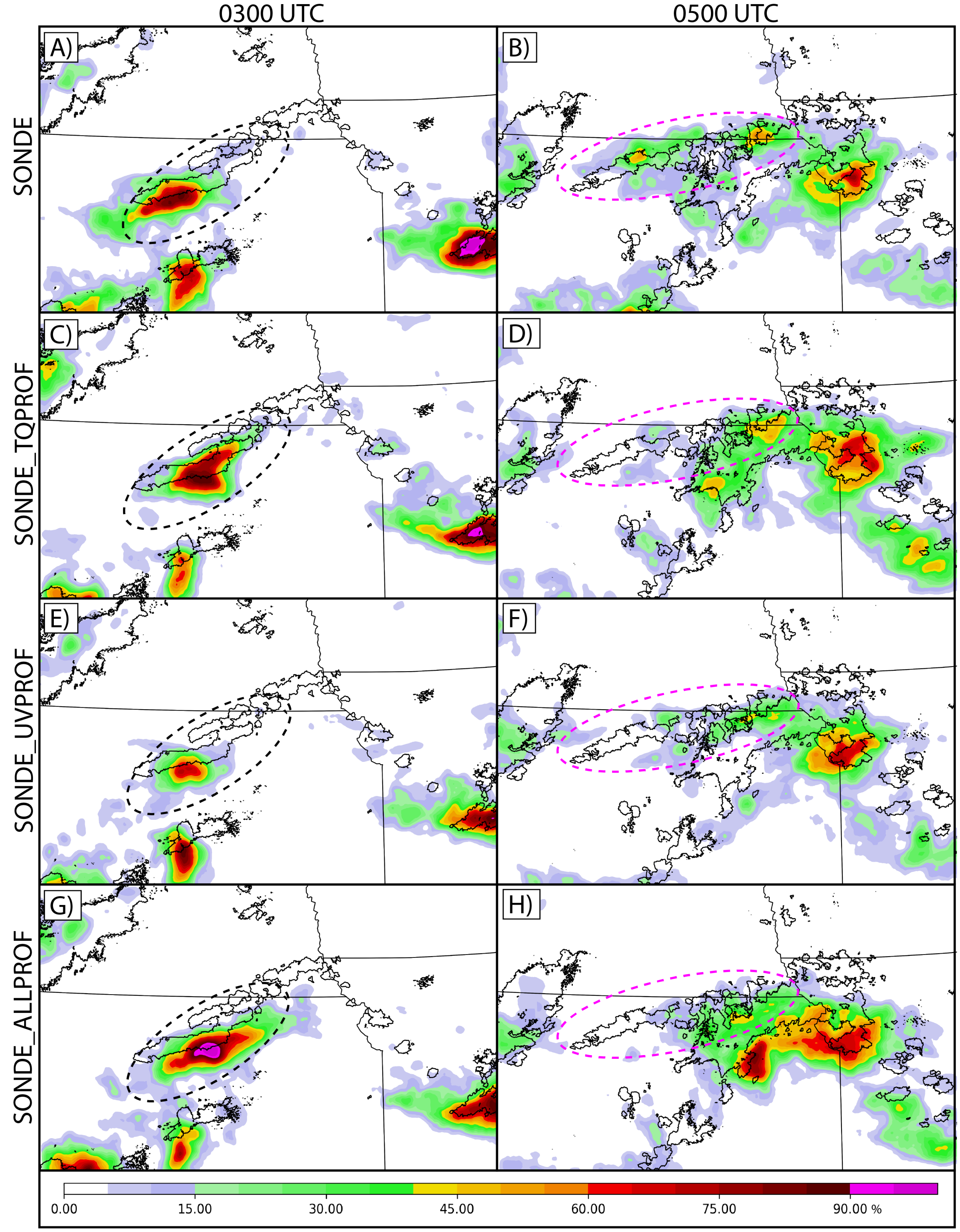


Figure 4.16: As in Fig. 4.12 but for the 26 June frontal overrunning case valid at (a,c,e,g) 0300 UTC; and (b,d,f,h) 0500 UTC.

their observation errors (Fig. 4.11b), such data are thus likely responsible for the degraded forecast quality.

4.7 Discussion

This chapter evaluates the systematic impact of assimilating a network of ground-based remote sensing instruments alongside collocated rawinsondes for 13 nocturnal CI events. The PECAN project provides a valuable dataset for determining these impacts given the diverse forcing mechanisms and environments for each case. Additionally, as most nocturnal convection is forced by mechanisms above the surface, such events serve as useful examples to understand the benefit of profiling instruments specifically. Although the limited number of nocturnal CI events sampled during PECAN prevents the estimation of statistical significance of the observation impacts, the results shown here offer valuable information compared to previous case studies that only assimilate data for one case or from a single observing site.

We find small but consistent improvements across nearly all verification metrics when assimilating thermodynamic profilers alongside collocated rawinsondes. These improvements include an increase in mean FSS in addition to improvements to standard contingency metrics such as POD, CSI, and bias. We find that assimilating the thermodynamic profiling data primarily results in midlevel cooling and moistening prior to CI, thus enhancing the likelihood of convective development and improving the detection of observed events. These results provide further evidence that a network of thermodynamic remote sensing profilers can have positive impacts on convective-scale forecasts.

Conversely, assimilating the collocated kinematic profilers often degrades the forecast performance for the detection of nocturnal CI, likely due to opposite impacts to the pre-convective thermodynamic fields compared to the direct observations from the AERIs. When both datasets are assimilated together, these detriments often counteract the benefits of assimilating the thermodynamic data. However, while assimilating the kinematic data

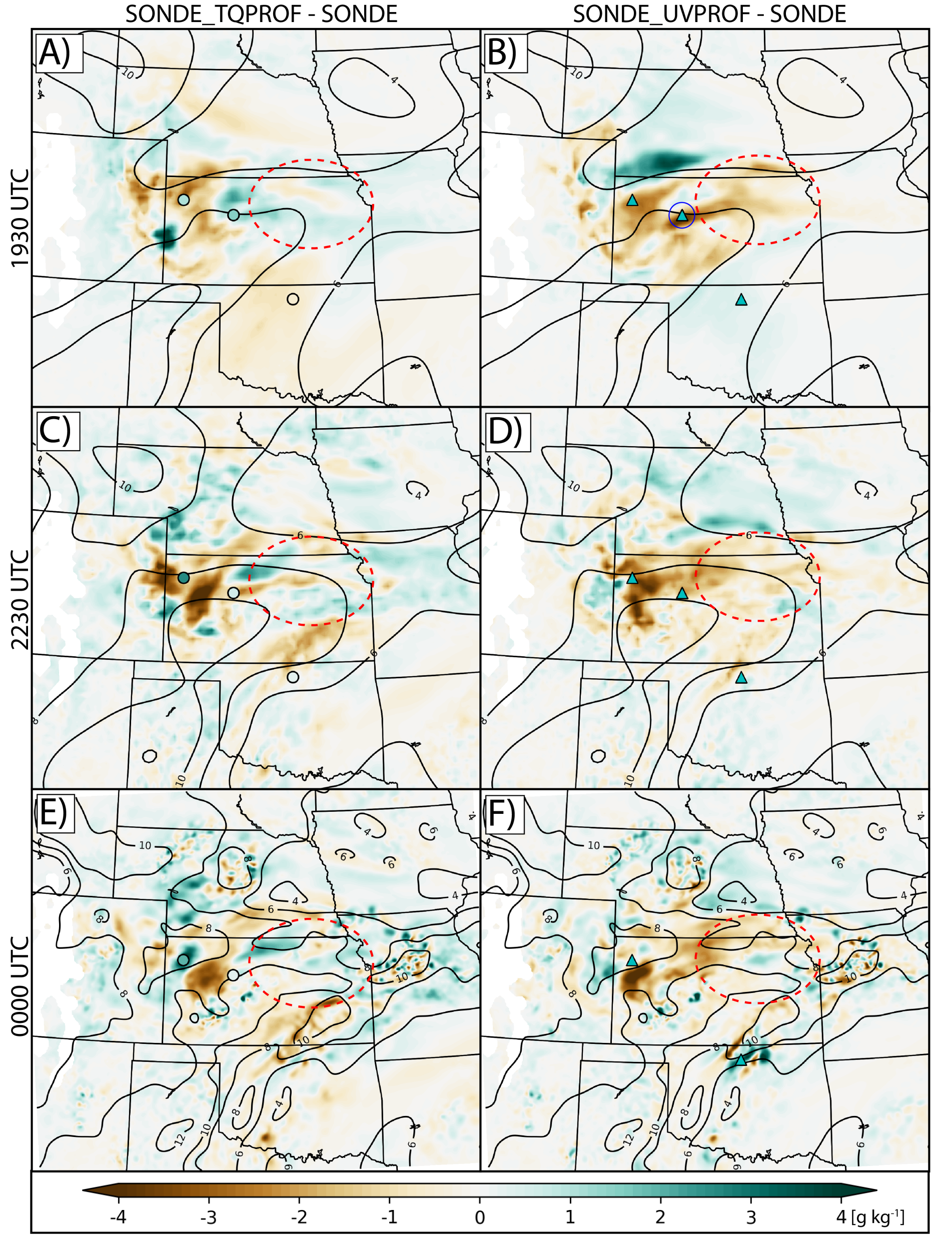


Figure 4.17: As in Fig. 13 but for 750 hPa water vapor mixing ratio ($g\ kg^{-1}$) and valid at (a,b) 1930 UTC 25 June; (c,d) 2230 UTC 25 June; and (e,f) 0000 UTC 26 June. The dashed red ovals correspond to the primary location of the frontal overrunning CI event for 26 June. The blue circle in (b) indicates the location of the FP3 profile shown in Fig. 4.18.

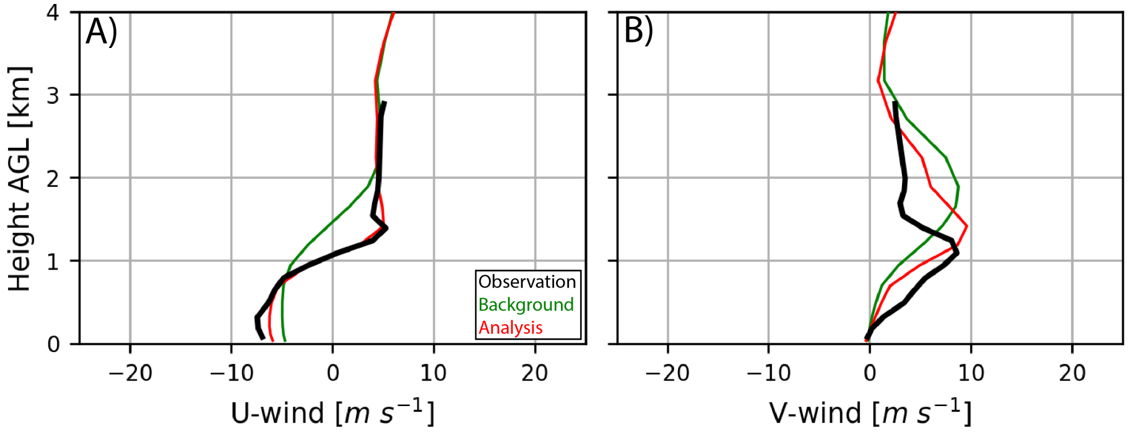


Figure 4.18: Ensemble mean background (green) and analysis (red) wind profiles for the 1930 UTC composite kinematic observation assimilated from FP3 (black; location circled in blue in Fig. 4.17b).

often degrades the detection of nocturnal CI, we find some forecast improvements when assimilating both datasets together. For example, the CI events that are successfully detected in SONDE_ALLPROF show lower timing, location, and orientation errors. To further improve the precise timing and location of CI, denser observing networks would likely need to be assimilated compared to the 2-6 sites assimilated here. Assimilating such a wider network of profilers would likely better constrain the location of convergence boundaries for which we saw little impact in this chapter. Finally, we also find strong variability in the impact of assimilating these datasets for different cases, likely due to the diverse mechanisms responsible for each nocturnal CI event and the number and relative location of each observing site.

We hypothesize that the forecast degradations from assimilating kinematic data are related in part to underestimated observation errors at some sites. These findings, along with the relationship between error inflation and forecast skill discussed in Chapter 3, suggest a large sensitivity between observation impacts and their assigned observation errors when assimilating high-frequency, remote sensing data. This sensitivity is likely further compounded by the large number of profilers that are assimilated at one time in this chapter (as many as 12 profiles from the same site during one DA cycle). As representation errors vary by environment, empirical error inflation methods such as the one applied here are likely not valid for a large set of cases. Thus, the next chapter of this dissertation examines flow-dependent methods to diagnose observation error variances for high-frequency profilers, including adaptively inflating observation errors (e.g., Minamide and Zhang 2017) or using background and analysis-error statistics to derive optimal observation errors (e.g., Desroziers et al. 2005).

Chapter 5

Convective-scale forecast impacts of flow-dependent and static observation error estimation for the frequent assimilation of thermodynamic profilers

5.1 Introduction

As discussed in Chapter 1, DA involves computing the posterior probability density function of a model state given prior probabilities and observation likelihoods (Kalnay 2003). Much previous DA research focuses on improving the structure of the former through ensemble- or hybrid-based methods that can diagnose flow-dependent background error correlations (Whitaker and Hamill 2002; Wang et al. 2007, 2008a,b, 2009, 2013; Buehner et al. 2010; Johnson et al. 2015). Conversely, few studies explore potential improvements to the estimate of the observation error covariance, **R**, despite the background and observation errors playing an equally important role. Much of this research gap is due to the prior assumption that conventional observations feature errors that change little in time or space and thus can be represented by a constant, Gaussian-distributed error variance (Fowler and Van Leeuwen 2013). However, novel remote sensing instruments have become increasingly common for assimilation in operational systems across the globe (e.g., Geer et al. 2018). Such instruments can be highly beneficial in improving NWP forecasts (Hu et al. 2019; Degelia et al. 2019; Chipilski et al. 2020) but are also known to feature highly variable observation errors. As such, it has become clear that adaptive and flow-dependent methods for diagnosing **R** should be further explored (Minamide and Zhang 2017; Fielding and Janiskova 2018; Fielding and Stiller 2019).

Observation errors are primarily composed of three main sources: instrument errors, systematic errors, and representation errors (Fowler and Van Leeuwen 2013). Instrument errors represent random observation noise that is typically small and can be represented by a constant error. Systematic errors are those that occur when observations are increasingly inaccurate in certain scenarios, such as when measuring a specific feature or at a different time of day. For example, some remote sensing profilers are known to perform poorly in the presence of clouds or temperature inversions (e.g., Turner and Löhnert 2014). Finally, representation errors occur due to mismatches between the observed and modeled variable. These can include errors associated with the observation operator that occur when mapping the model state into the observed state (e.g., errors in a radiative transfer model). Representation errors can also occur due to mismatched scales between the observation and model state, such as when a coarse model grid box smooths through a frontal boundary or shallow inversion. Given their relation to the atmospheric state, systematic and representation errors are sometimes referred to as "flow-dependent" observation errors (Minamide and Zhang 2017).

Many datasets are known to feature heavily flow-dependent observation errors. For example, Fielding and Janiskova (2018) show that satellite-based, radar observation errors increase from $\sim$1 dB in the stratiform region to $\sim$8 dB within the convective core of a simululated MCS. Despite the strong variation in these errors, most operational DA systems assume constant (static) observation errors that do not vary in time or location (Bormann et al. 2016). These assumptions naturally result in some observations being underweighted (overestimated observation errors) or overweighted (underestimated observation errors) during DA. We document such problems in Chapter 4 when assimilating ground-based remote sensing data collected during PECAN. When using a static error inflation method across the entire network of profilers, we find that some profiling sites feature underestimated errors. Assimilating data from these sites are then shown to weaken midlevel moisture advection and degrade subsequent forecasts of nocturnal CI.

To resolve these problems, a few recent studies propose methods for diagnosing flow-dependent observation errors in DA. Geer and Bauer (2011) adaptively inflate static observation errors for cloud-top observations based on an empirical function of brightness temperature. Minamide and Zhang (2017) use a similar method but only apply the error inflation when the difference between the observed and simulated brightness temperature (i.e., the innovation) is large. Both methods reduce analysis errors by limiting the use of large observation increments due to mismatched scales. Instead of relating representation errors to an empirical function, Fielding and Stiller (2019) develop a method for directly computing representation errors based on the local variance of an observation and a climatological correlation factor. Though they do not assimilate observations using these methods, they show that such methods can accurately estimate representation errors routinely observed by satellite-based radar instruments. Finally, a variety of studies also employ a method derived in Desroziers et al. (2005, hereafter "D05") for computing the full observation error matrix (i.e., all three error components; Weston et al. 2014; Bormann et al. 2016; Cambell et al. 2017). The D05 method uses observation-space diagnostics, including the innovation and analysis residuals, to compute an estimate of $\mathbf{R}$.

While these previous studies present some advantages of assigning flow-dependent observation errors, many questions are still yet to be addressed. First, the methods described in Geer and Bauer (2011) and Minamide and Zhang (2017) are designed to inflate errors only due to representation errors resulting from mismatched scales. This ignores potential systematic errors that occur when instruments are less accurate in certain conditions. Additionally, previous applications of the D05 method typically compute a statistically robust estimate of $\mathbf{R}$ by collecting innovation and analysis statistics from many DA cycles (e.g., Weston et al. 2014; Bormann et al. 2016; Cambell et al. 2017). These applications effectively average the observation errors and ignore potential variations that can occur in some meteorological conditions. Finally, apart from Minamide and Zhang (2017), most studies described above only evaluate the impact of observation errors on synoptic or coarser mesoscale models. As such, no known work has determined the convective-scale forecast impact of utilizing flow-dependent observation errors. To address these research gaps, we propose two methods for estimating observation errors that account for both time-varying systematic and representation errors. Specifically, we compare static and flow-dependent methods to evaluate whether such novel methods can improve forecasts of a nocturnal MCS observed during PECAN on 15 July 2015. We evaluate these methods by assimilating data collected by AERIs, as these instruments are known to feature variable performance statistics depending on the meteorological conditions (Turner and Löhnert 2014; Turner and Blumberg 2019).

The outline of this chapter is as follows: section 5.2 presents an overview of the methods used to assign the observation errors including their static formulation and a flow-dependent extension. System configuration and details of the 15 July nocturnal MCS used to evaluate these methods are discussed in section 5.3. Attributes of the error profiles diagnosed by each method are presented in section 5.4. The impacts of assimilating these error profiles for the convective event are detailed in section 5.5, and the impacts on the DA cycling are discussed in section 5.6. Finally, a discussion of results is found in section 5.7.

5.2 Methods for assigning observation errors: static methods and their flow-dependent counterparts

This study compares the convective-scale impact of assimilating AERIs using both static and novel, flow-dependent methods to assign observation errors. These remote sensing instruments retrieve simultaneous profiles of temperature and water vapor mixing ratio up to 3 km AGL every 5-15 minutes. As the accuracy of AERI retrievals is sensitive to various atmospheric parameters including clouds and precipitable water vapor (Turner and Löhnert 2014; Turner and Blumberg 2019), observation errors for these data are likely highly flow-dependent and therefore act as a useful tool to evaluate various observation estimation methods. The following is a summary of the methods evaluated in this study including descriptions of both their static and flow-dependent implementations. We note that the serial EnKF applied here (described in section 2.1) assumes a diagonal **R** with no observation error correlations. As such, the methods below are only designed to compute observation error variances instead of covariances.

5.2.1 Inflation of AERIoe retrieval errors

We originally introduce a method that assigns observation errors for AERI data by statically inflating the retrieval errors generated by AERIoe (refer to section 3.3.1 and Eq. 3.1). This method inflates the errors using the difference between the full error profiles used for assimilating rawinsondes in GSI and rawinsonde instrument errors. The weight of the error inflation is controlled by the tunable parameter α that is different for temperature and moisture and linearly increases with height. Though the initial retrieval errors are unique for each observing time, the additive inflation (second term on RHS of Eq. 3.1) tends to be much larger such that Eq. 3.1 produces a primarily static error profile with little variance between observing times. By comparing profiles of the final, inflated error to RMS differences (RMSD) from collocated rawinsondes, we find that this static inflation method produces a reasonable shape and magnitude for the mean error profiles but greatly underestimates the variability in the observation errors (Fig. 5.1). Upper-level moisture observations often differ from rawinsonde observations by 0.5-4 $g\ kg^{-1}$, while the observation errors diagnosed using Eq. 3.1 only vary by 1.8-2.2 $g\ kg^{-1}$. As such, the inflation method from section 3.3.1 can sometimes cause upper-level observations to be underweighted (overestimated errors) or overweighted (underestimated errors), likely leading to suboptimal observation increments.

To introduce flow-dependence into the AERIoe inflation method, we modify the method from section 3.3.1 to include the effective vertical resolution profile for the retrieval (γ; units of km). AERIoe outputs a unique profile of γ for each observing time that measures

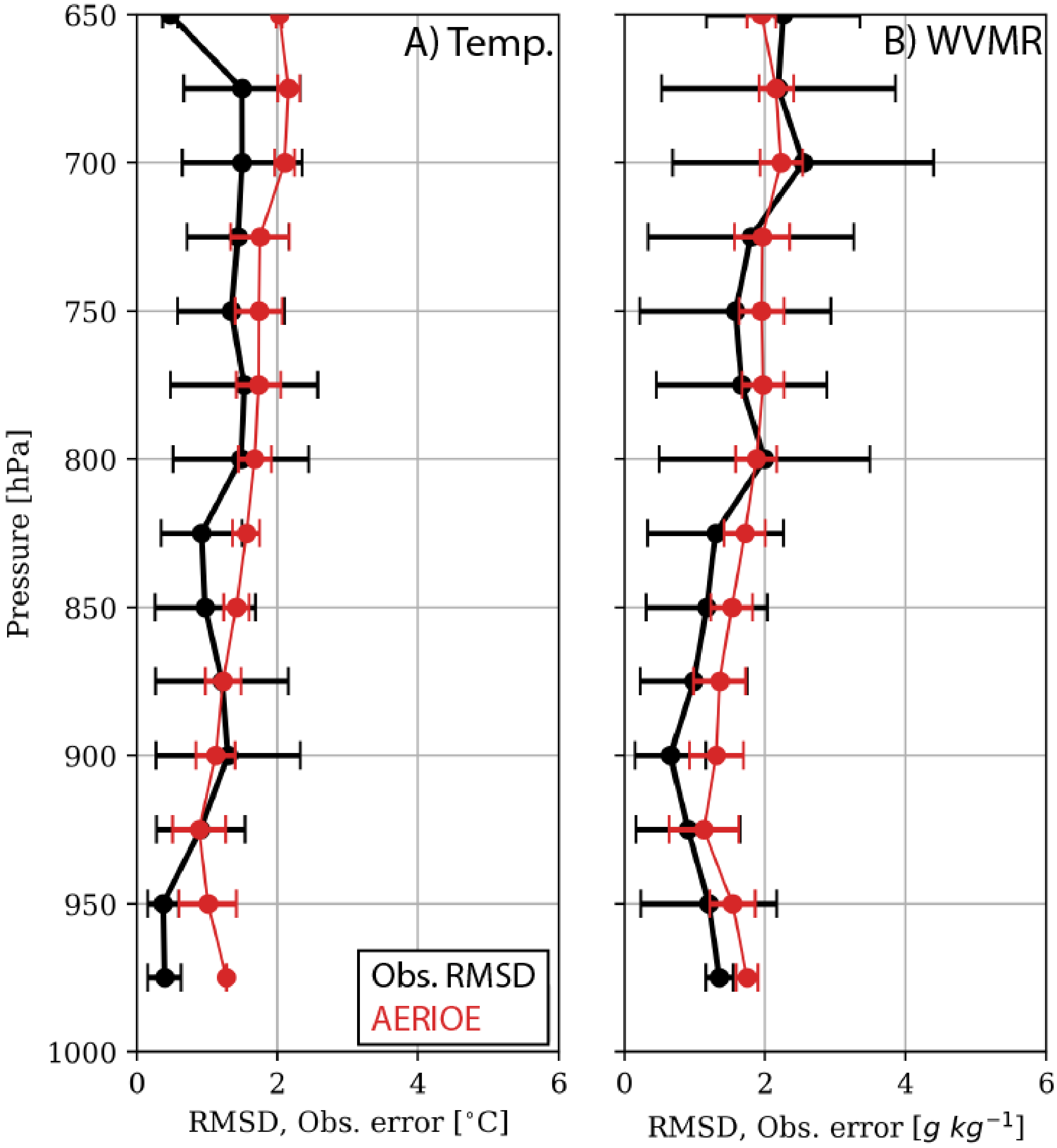


Figure 5.1: RMS differences between AERI retrievals and collocated rawinsonde launches for all retrievals assimilated in Chapter 4 (black). AERI retrievals are only shown if rawinsonde data is available at the same site within a ±15 min window. The error bars in black represent one standard deviation of the absolute differences between the retrievals and rawinsondes. The RMS error for rawinsondes has also been subtracted from the black curve. Also shown in red are the mean and standard deviation of the observation error profiles assigned for the same retrievals using Eq. 3.1. The mean differences are computed by interpolating each AERI retrieval and rawinsonde onto a standard vertical grid with 25 hPa spacing.

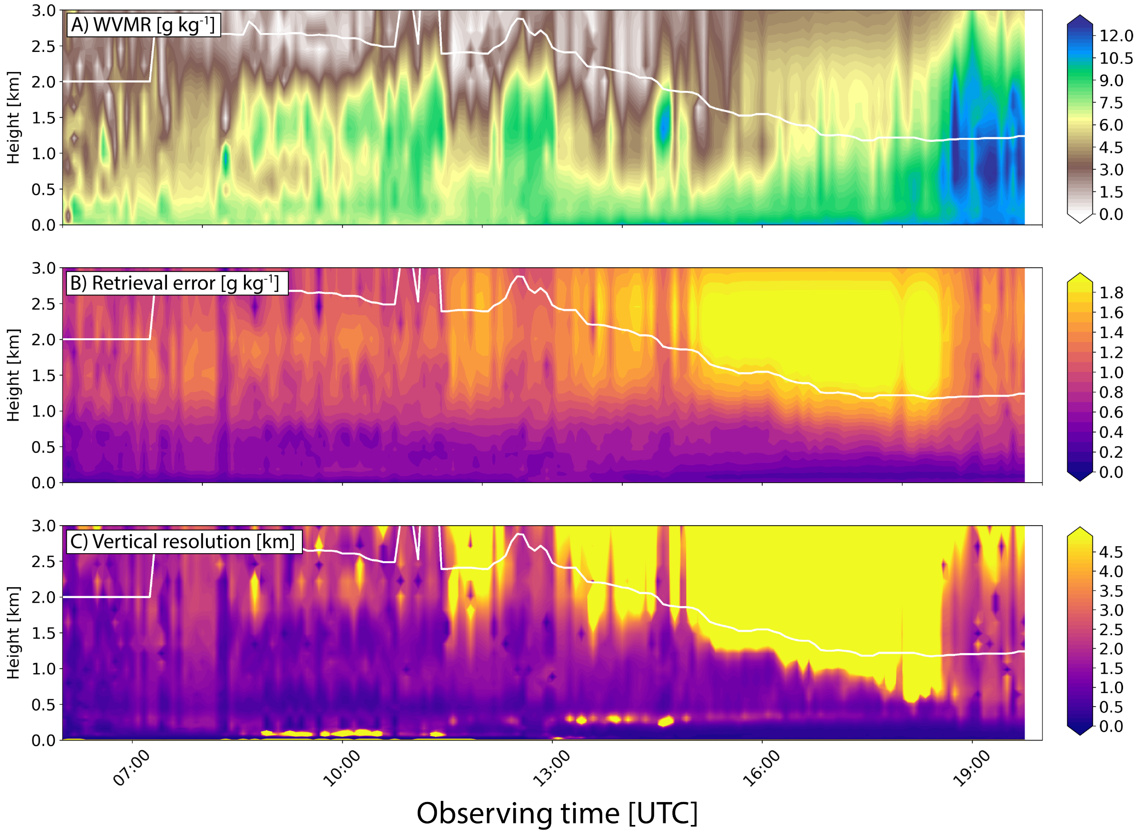


Figure 5.2: Example of AERIoe moisture retrievals collected between 0600-2000 UTC 23 June during PECAN. Included are the (a) moisture observations ($g\ kg^{-1}$), (b) retrieval errors ($g\ kg^{-1}$), and (c) retrieval effective resolution profiles (km). Also overlaid on each panel is the cloud base height (km) indicated by a collocated lidar or ceilometer (white).

the vertical smoothing used in the retrieval. We refer to Turner and Löhnert (2014) for the derivation of γ within AERIoe. Following this modification, Eq. 3.1 becomes:

$$\sigma_{Pf}^2 = \sigma_{Pi}^2 + \beta\gamma(\sigma_{Sf}^2 - \sigma_{Si}^2) \tag{5.1}$$

where β represents a similar tunable parameter to α but now has units km^{-1}. The AERIoe effective resolution typically increases with height (i.e., more smoothing at the top of the profile; Fig. 5.2a). This results in Eq. 5.1 producing a similar effect to Eq. 3.1 wherein the error inflation increases with height. As such, we do not require β to increase with height as with α in Eq. 3.1.

Whereas the method in Eq. 3.1 equally inflates each error profile, Eq. 5.1 instead causes the AERI retrieval errors to be adaptively inflated. For example, this method often inflates errors more in the shallow layer below cloud base, where Turner and Blumberg (2019) note that the AERI effective resolution becomes large due to problems separating cloud emission from atmospheric emission. An example of this issue is shown for a PECAN case in Fig. 5.2 where the AERI vertical resolution rapidly grows through a layer beginning ~500 m below cloud base after 1100 UTC. While the retrieval errors increase by a factor of 5 relative to the surface (Fig. 5.2b), the effective vertical resolution instead increases by a factor of 20 throughout the same layer (Fig. 5.2c). This indicates considerable smoothing in the observations that is not fully accounted for by the retrieval errors. Thus, applying the flow-dependent error inflation in Eq. 5.1 can correspondingly inflate the retrieval errors to account for such problems.

5.2.2 Desroziers et al. (2005) diagnostic

In addition to inflating errors produced by AERIoe, we also evaluate the impact of assigning error variances derived from observation-space diagnostics (Tandeo et al. 2018). The most common implementation for this follows from D05 who derive an estimate of $\mathbf{R}$ as a relationship between the innovation ($\mathbf{d_b^o} = \mathbf{y^o} - \mathbf{Hx^b}$) and analysis residuals ($\mathbf{d_a^o} = \mathbf{y^o} - \mathbf{Hx^a}$). Here, $\mathbf{y^o}$ is the observation vector, and $\mathbf{Hx^b}$ and $\mathbf{Hx^a}$ are the background and analysis vectors in observation space. Assuming that observation and background errors are uncorrelated and that the covariances used to compute $\mathbf{x^a}$ are consistent with the true matrices, D05 show that:

$$\widetilde{\mathbf{R}} = E[\mathbf{d_a^o}(\mathbf{d_b^o})^T] \tag{5.2}$$

where $\widetilde{\mathbf{R}}$ approximates the true $\mathbf{R}$. Because $\widetilde{\mathbf{R}}$ contains information about the model departure from the observation, it should ideally contain all components of the observation error including systematic errors and errors related to representation. We note that the D05 diagnostic is often referred to as a posterior "consistency check" on the assigned observation errors given its requirement for an existing analysis vector. To assimilate observations using the D05 diagnostic, Bormann et al. (2011) propose a method wherein an assumed set of errors are used to compute an initial analysis. $\widetilde{\mathbf{R}}$ is then diagnosed using Eq. 5.2 and the analysis update is repeated using the new observation error covariance matrix. This analysis-diagnostic cycle can also be iterated to improve the estimate of $\widetilde{\mathbf{R}}$.

Recently, Cambell et al. (2017) show that remote sensing instruments often feature large, correlated observation errors. For such instruments, the ability of the D05 method to diagnose a full error covariance matrix, including off-diagonal terms related to error correlations, could be particularly useful. Nevertheless, many operational DA systems require a diagonal $\mathbf{R}$ under the assumption that conventional datasets do not typically feature significant observation error correlations. Previous studies show that ignoring the off-diagonal terms of $\widetilde{\mathbf{R}}$ by only assigning observation error variances (i.e., the diagonal terms of $\widetilde{\mathbf{R}}$) can lead to sub-optimal analyses (Bormann et al. 2016). This problem can be at least partially alleviated by applying a multiplicative inflation to the error variances (Stewart et al. 2013; Bormann et al. 2016). Following these findings, we use the D05 diagnostic for assigning AERI observation errors through:

$$\sigma_{Pf}^2 = C diag(\widetilde{\mathbf{R}}) \tag{5.3}$$

where C is a constant multiplicative inflation factor.

Previous studies typically compute $\widetilde{\mathbf{R}}$ using a large subset of observations to increase the statistical robustness of the estimated errors. For example, Bormann et al. (2016) compute $\widetilde{\mathbf{R}}$ from a collection of satellite data assimilated over one month. Given that a large observation subset essentially averages the observation errors over an extended period, this standard implementation of Eq. 5.3 produces static errors that are not a function of the background state. However, previous research supports potentially reducing the window used for the D05 diagnostic, as Weston et al. (2014) find that $\widetilde{\mathbf{R}}$ for sounder observations converges on a solution within one day. Cambell et al. (2017) also note finding little daily variability

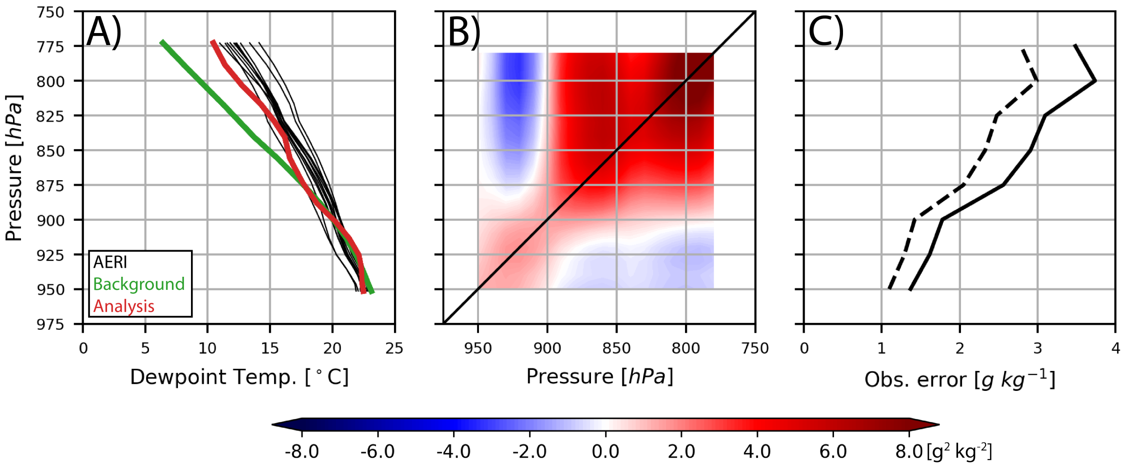


Figure 5.3: (a) AERI moisture retrievals (dewpoint temperature; °) and corresponding background and analysis profiles from FP5 and assimilated during the 0400 UTC 15 July DA cycle. Also shown are (b) the observation error covariance matrix ($\widetilde{\mathbf{R}}$; $g^2\ kg^{-2}$) diagnosed by Eq. 5.2 for the same retrievals; and (c) the uninflated and inflated diagonal of $\widetilde{\mathbf{R}}$ assimilated in the DESROZIERS_FD experiment ($g\ kg^{-1}$; Eq. 5.3). In (c) Eq. 5.3 is computed using $C = 1.25$. We note that the matrix in (b) is not symmetric given that $\widetilde{\mathbf{R}}$ is only an estimate of the true $\mathbf{R}$.

in the computed diagnostics. Following these findings, we also evaluate a flow-dependent implementation of D05 wherein a unique $\widetilde{\mathbf{R}}$ is computed during each DA cycle. This method effectively sacrifices statistical robustness for time-dependent observation errors. As shown in Fig. 5.3, the error profiles diagnosed by the flow-dependent implementation of D05 are primarily related to the magnitude of the innovation and analysis residuals such that the errors increase near the top of the profile due to the AERI retrievals differing from the model state.

5.2.3 Experimental design and implementation of error methods

To evaluate the convective-scale impact of these methods, we design a set of experiments that differ only in their methods used to assign the observation errors when assimilating AERI retrievals (Table 5.1). First, we perform a reference experiment where no AERI data are assimilated (NOAERI) to ensure that the latter experiments do not overinflate the observation errors and result in little impact on the analysis. Next, we perform two experiments that inflate the AERIoe retrieval errors, including a method based on the static error inflation from Eq. 3.1 (AERIOE), and another that uses the flow-dependent extension in Eq. 5.1 wherein the errors are adaptively inflated (AERIOE_FD). Additionally, this study performs one of the first known evaluations of applying the D05 observation-space diagnostics for convective-scale DA. Two experiments are included that use this method, including a static but statistically robust implementation where Eq. 5.3 is computed *a priori* using the 13 cases evaluated in Chapter 4 (DESROZIERS). For this experiment, separate error profiles are computed for each assimilation domain given that representation errors are highly related to model grid spacing (Fielding and Stiller 2019). We also evaluate a flow-dependent extension of the D05 method where Eq. 5.3 is computed "in-line" during the assimilation cycling (DESROZIERS_FD). For this experiment, we apply the iterative method for computing $\widetilde{\mathbf{R}}$ (Bormann et al. 2011) and find that the error profiles converge within three iterations. We also note that separate error profiles are computed for each AERI observing site in both DESROZIERS and DESROZIERS_FD.

The two flow-dependent experiments (AERIOE_FD, DESROZIERS_FD) are likely more impactful in certain situations depending on the atmospheric state. For example, we hypothesize that AERIOE_FD can improve analysis results by inflating observation errors below cloud base where systematic errors often occur. DESROZIERS_FD could instead limit observation impacts in situations where the retrieval smooths through an inversion or shallow moist layer, resulting in a large innovation. To combine the benefits of these two methods, we perform an additional experiment (MAX) that assigns observation errors based on the maximum value diagnosed by either AERIOE_FD or DESROZIERS_FD:

$$\sigma_{Pf}^2 = max[\sigma_{Pf}^2(AERIOE\_FD), \sigma_{Pf}^2(DESROZIERS\_FD)] \tag{5.4}$$

MAX follows a similar approach to Minamide and Zhang (2017) who select the maximum error between a baseline observation error (instrument error plus a constant representation error) and an error estimate derived from observation-space diagnostics similar to Eq. 5.2.

Table 5.1: List of experiments in Chapter 5, a description of the method used to assign their observation errors, and values for the tunable parameters in each equation

Experiment	Description of method	Parameter values
NOAERI	N/A, no AERI observations assimilated	N/A
AERIOE	Initial observation errors obtained from AERIoe retrieval and statically inflated following Eq 3.1.	$\alpha_T = [0, 4\ at\ 3\ km\ AGL]$ $\alpha_Q = [0, 1\ at\ 3\ km\ AGL]$
AERIOE_FD	Initial observation errors obtained from AERIoe retrieval and adaptively inflated using retrieval effective resolution in Eq. 5.1.	$\beta_T = 1.3\ km^{-1}$ $\beta_Q = 0.2\ km^{-1}$
DESROZIERS	Static observation errors obtained using observation-space diagnostics in Eq. 5.3. Errors are computed a-priori using all DA cycles simulated in D20.	$C = 1.25$
DESROZIERS_FD	Flow-dependent observation errors obtained using observation-space diagnostics in Eq. 5.3. Errors are uniquely computed during each DA cycle.	$C = 1.25$
MAX	Maximum observation error between AERIOE_FD and DESROZIERS_FD	Refer to AERIOE_FD and DESROZIERS_FD

The MAX experiment is also designed to alleviate instances where DESROZIERS_FD produces an observation error that is too low, such as when the innovation or analysis residuals are near zero.

Each assimilation method described above uses a tunable parameter (α, β, C) to control the magnitude of the observation errors (values listed in Table 5.1). To be consistent with previous results, we use the values of α applied in section 3.3.1 for the AERIOE experiment (see Fig. 3.3). For the parameter β used in AERIOE_FD, we select values that result in the lowest mean absolute difference between observation errors produced by AERIOE and AERIOE_FD. As such, this flow-dependent modification increases the profile-to-profile variation of the errors in AERIOE_FD while only slightly changing the mean errors (primarily at the top of the profile). Finally, to be consistent with the other experiments, we also tune C using the 26 June 2015 case that is used to select α in Chapter 3. We find that C=1.25 produces the highest FSS and thus we also use that value to simulate the 15 July MCS presented herein.

5.3 Case description and system configuration

5.3.1 Overview of the 15 July 2015 MCS

We evaluate each experiment for the nocturnal MCS observed during PECAN on 15 July 2015 (IOP30). Grasmick et al. (2018) provide a detailed overview of this event which we summarize here. On the late afternoon of 14 July, a cluster of disorganized convective cells developed in eastern Colorado along a group of shortwave troughs embedded within an upper-level ridge. Additional convection initiated along outflow boundaries produced by these storms shortly after sunset, eventually growing upscale into the MCS of interest (Fig. 5.4a). By 0700 UTC, this MCS developed the traditional leading-line, trailing stratiform structure with the strongest precipitation located along the northern half of the MCS (Fig. 5.4b). An additional band of convection, oriented primarily east-west, developed just north of the main convective line near the Nebraska border at 0800 UTC (Fig. 5.4c). This secondary convective line will be discussed throughout this chapter. The main MCS continued to propagate eastward, maintaining its structure before eventually decaying around 1200 UTC (Fig. 5.4d). Though not considered a severe MCS (Grasmick et al. 2018), the NWS reported three heavy wind events (30 ms^{-1} or 60 kt) and one large hail event during its evolution.

We previously assimilate both thermodynamic and kinematic profilers for this case as part of the systematic study detailed in Chapter 4. While the baseline forecasts perform well, we find that assimilating the AERI data results in reduced skill for accumulated precipitation forecasts (not shown). The forecasts are most degraded between 0700-0900 UTC during which the MCS is heavily organized and when the secondary, east-west convective band develops to the north (Fig. 5.4c). As discussed in section 4.7, these forecast degradations could partially be a result of using static observation errors for datasets that feature large,

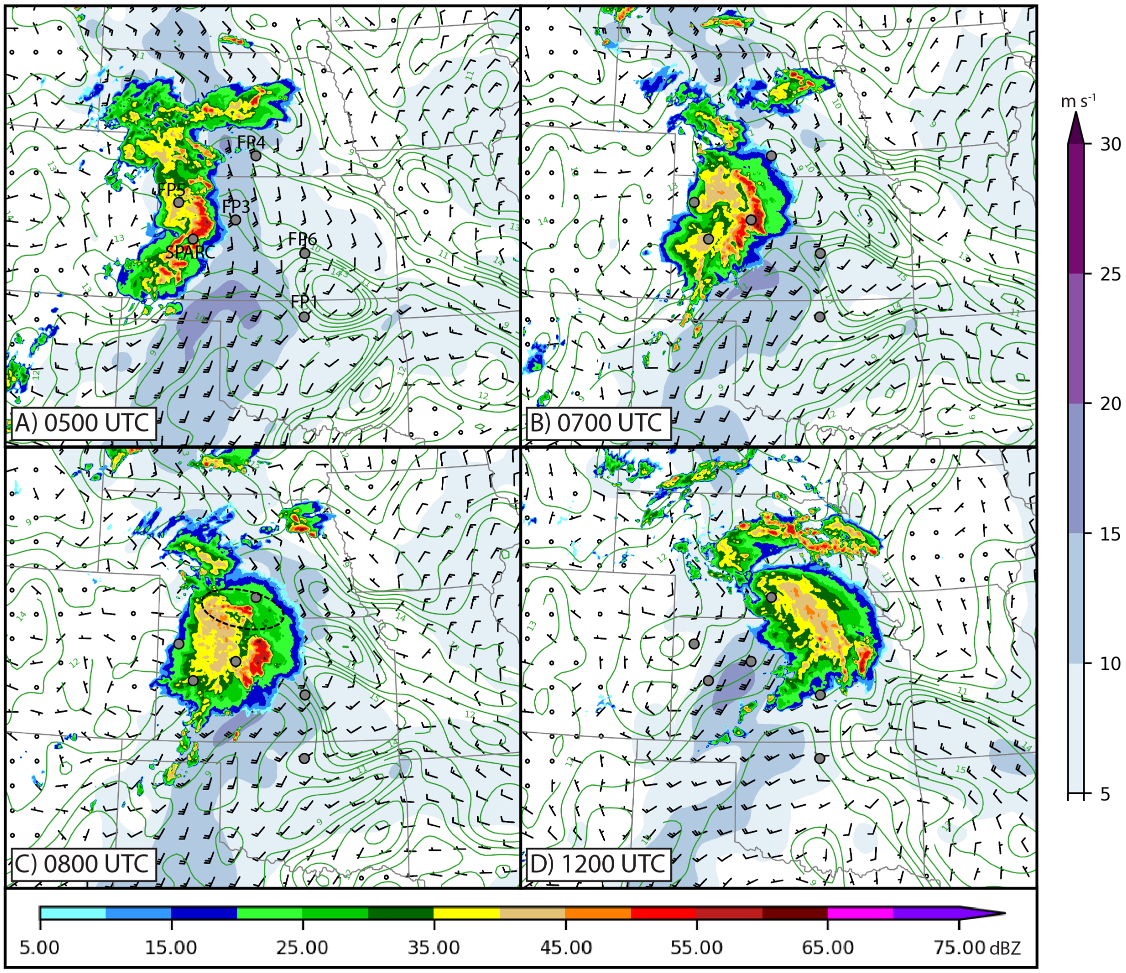


Figure 5.4: RAP analyses for the 15 July 2015 nocturnal MCS event including composite reflectivity (dBZ; bottom color scale), 850 hPa winds (ms^{-1}; right color scale), and 850 hPa water vapor mixing ratio ($g\ kg^{-1}$; green contours). The location of the AERI platforms assimilated here are overlaid in each panel and labeled in (a).

Table 5.2: As in Table 2.3 but only for the AERI instruments assimilated in Chapter 5.

Site name	Location	Instruments	Reference
FP1	Lamont, OK	AERI (AERIprof)	Gero et al. 2014
FP3	Ellis, KS	AERI (AERIoe)	Turner (2016a)
FP4	Minden, NE	AERI (AERIoe)	Turner (2016b)
FP5	Brewster, KS	AERI (AERIoe)	Turner (2016c)
FP6	Hesston, KS	AERI (AERIoe)	Turner (2016d)
MP3 "SPARC"	Mobile	AERI (AERIoe)	Wagner et al. (2016b)

flow-dependent error components. Given these findings, the 15 July MCS is selected to evaluate the impact of these novel methods. Additionally, we note that no kinematic profilers are assimilated here given that the methods discussed in section 5.2 are primarily designed for AERIs (see assimilated sites in Table 5.2 and Fig. 5.4).

5.3.2 Configuration of DA cycling

We evaluate the experiments described above using the same multi-scale DA and forecast system described in Chapter 2. The cycled DA begins by generating initial and lateral boundary conditions at 1600 UTC 14 July. Next, conventional data and AERI retrievals are assimilated on the 12-km outer domain (d01 in Fig. 5.5) at 3-h intervals until 0400 UTC 15 July. After DA on d01, we then downscale to the inner, 4-km domain (d02 in Fig. 5.5) during which conventional data, AERI retrievals, and radar observations are assimilated at 15-min intervals until 0530 UTC 15 July. Finally, we initialize 7-h forecasts from the first 10 ensemble members following Chapter 4. However, unlike Chapter 4 where we analyze small-scale nocturnal CI events on a 1-km grid, this study focuses on the larger-scale MCS that spans from southern Nebraska to northern Oklahoma. As such, we only analyze forecasts on the 4-km domain to limit computational costs.

5.4 Diagnosed observation error profiles

To examine the general properties of the observation errors diagnosed by each method, we first present mean error profiles computed over all DA cycles (Figs. 5.6-5.7). The error

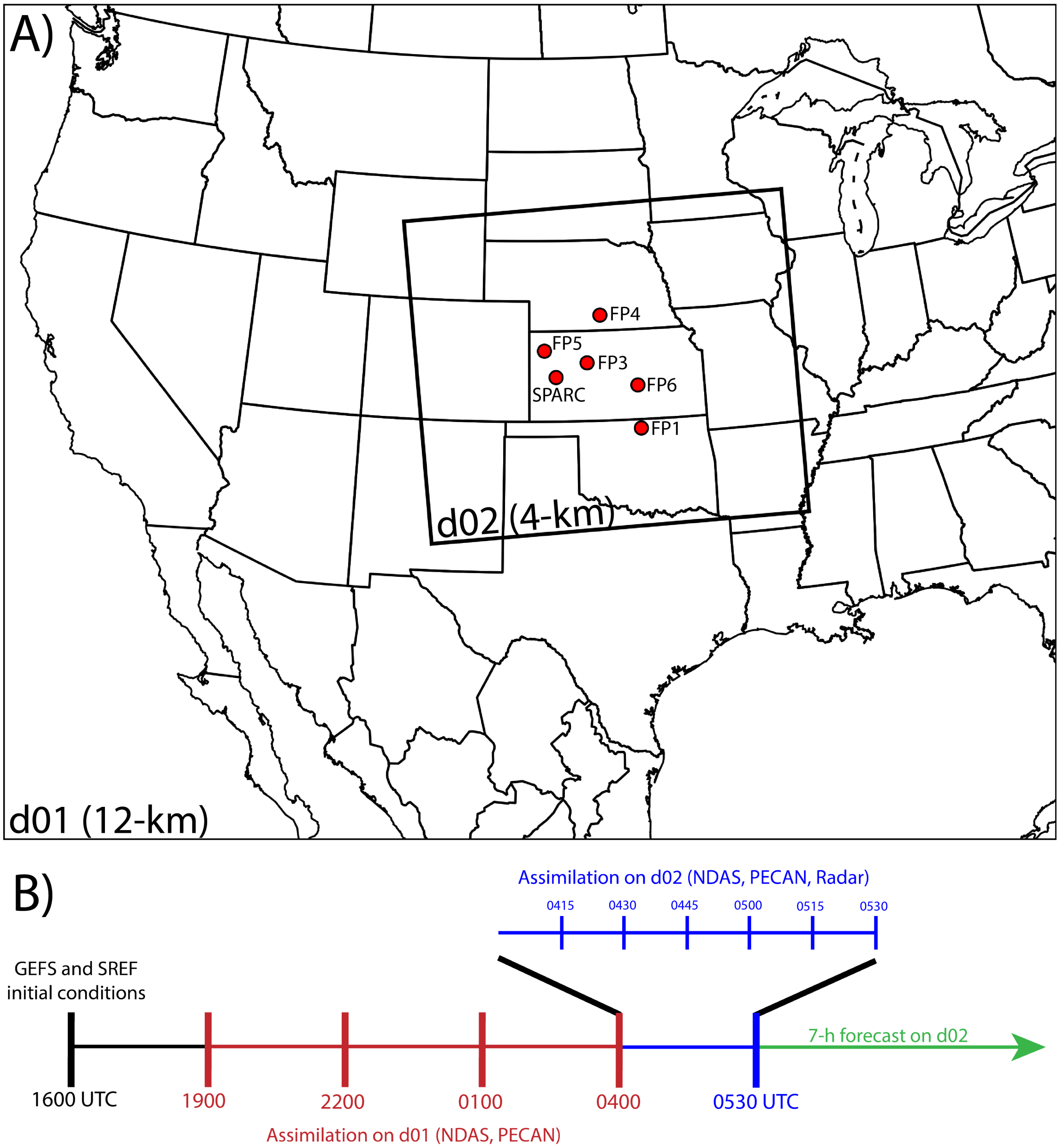


Figure 5.5: Overview of experimental design for Chapter 5 including (a) domain configuration for the outer (d01, 12-km) and inner (d02, 4-km) assimilation domains. The inner assimilation domain is also used as the forecast domain. The location of each AERI platform assimilated here are also overlaid. (b) Flowchart for the cycled assimilation, including the four 3-h assimilation cycles on d01 beginning at 1600 UTC 14 July, the six 15-min assimilation cycles on d02 beginning at 0415 UTC 15 July, and the 7-h forecast period beginning at 0530 UTC 15 July.

bars represent the standard deviation of the diagnosed errors and should not be confused with the "error standard deviation" used to assign observation errors within the DA system. Additionally, we choose not to present MAX here given that its observation errors are exactly equal to the larger errors between AERIOE_FD and DESROZIERS_FD.

In general, each method diagnoses similar error magnitudes for midlevel (∼900-750 hPa) temperature and moisture observations with larger differences appearing at the bottom and top of the retrievals. These differences primarily occur between the two main experiment sets (i.e., AERIOE and AERIOE_FD versus DESROZIERS and DESROZIERS_FD). When averaged across all sites, the methods which inflate the AERIoe retrieval errors produce vertically increasing errors up to 700 hPa for both temperature (Fig. 5.6a) and moisture (Fig. 5.7a). The DESROZIERS and DESROZIERS_FD methods instead show a local maximum in errors near the surface that decreases through the PBL. This result indicates that the innovation and analysis residuals are largest near the surface, potentially due to local circulations sampled by AERIs that cannot be resolved in the 12-km and 4-km simulations (i.e., larger representation errors). Differences between the two experiment groups are magnified at certain sites and especially for moisture errors. For example, DESROZIERS and DESROZIERS_FD indicate much larger errors in the midlevels at FP4, FP5, and FP6 (Fig. 5.7c,d,e), whereas the AERIoe inflation methods produce larger errors at SPARC (Fig. 5.7f). Overall, the DESROZIERS and DESROZIERS_FD methods also produce much more site-to-site variability due to their additional sensitivity to the background and analysis states.

The flow-dependent extensions of each method follow generally similar shapes to their static counterpart. AERIOE and AERIOE_FD produce similar error profiles below 750 hPa due to β in Eq. 5.1 being tuned to produce the lowest RMSD between the two experiments. However, the AERIOE_FD errors become much larger than AERIOE in the upper portions of the profile, especially for temperature (Fig. 5.6). This increase partially results from the effective vertical resolution term in Eq. 5.1 often increasing exponentially with height

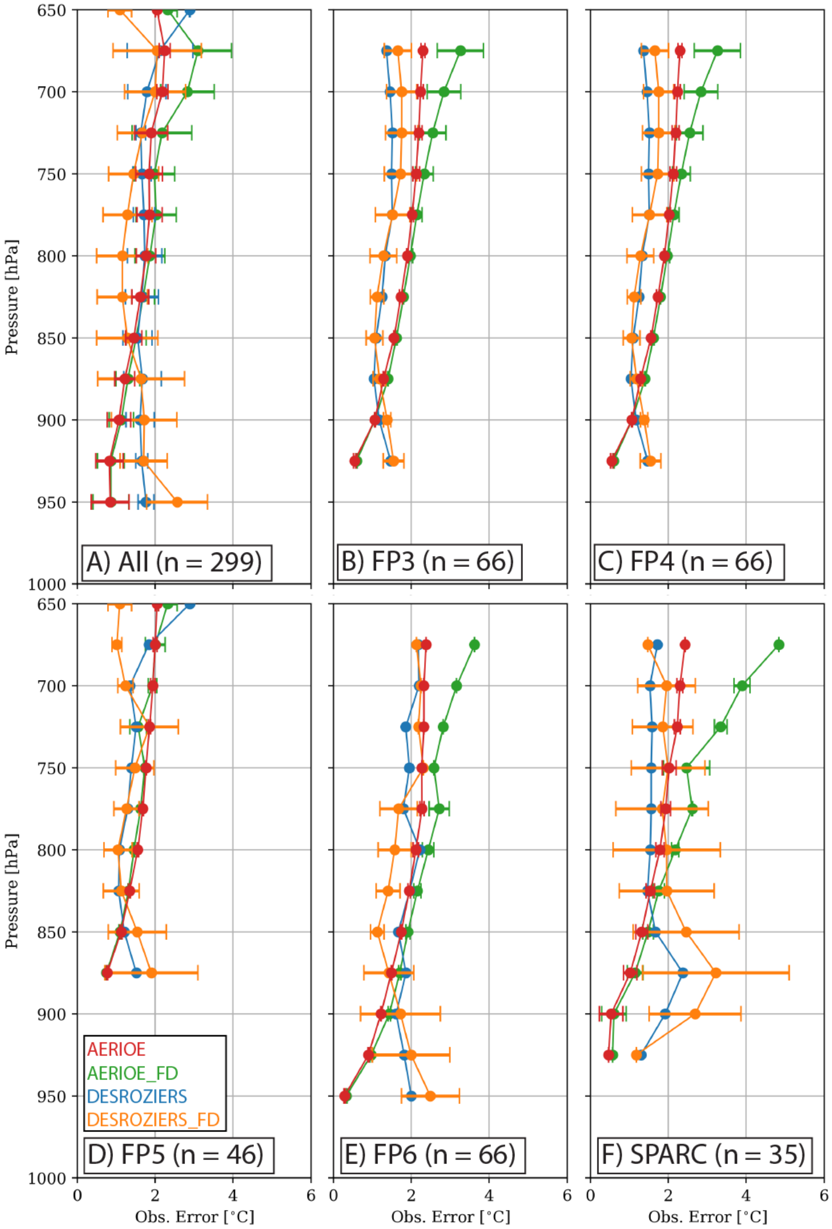


Figure 5.6: Mean and standard deviations of the observations error profiles for temperature (°C) from each site assimilated in Chapter 5 (except for FP1). The profile in (a) is computed using observation error profiles from all sites. When computing the statistics, each profile is interpolated onto a standard vertical grid with 25 hPa spacing.

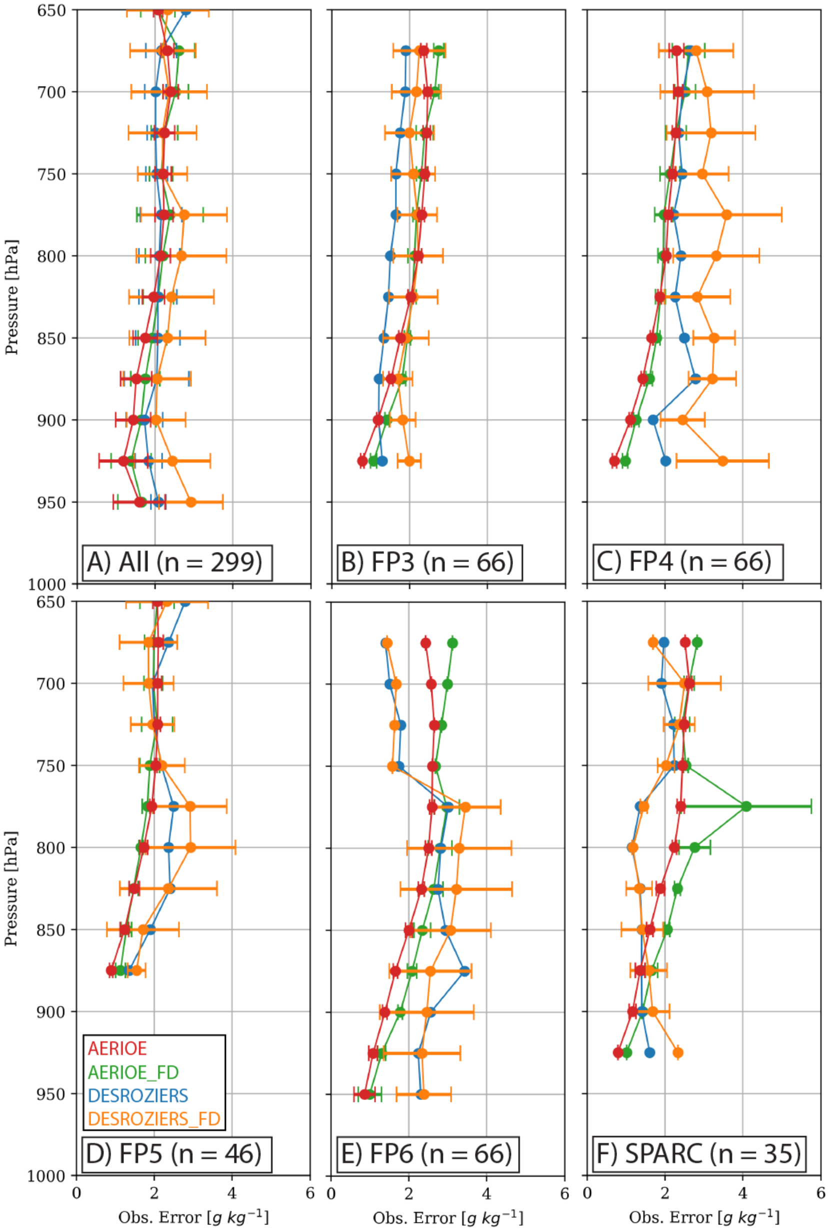


Figure 5.7: As in Fig. 5.6 but for water vapor mixing ratio observation errors ($g\ kg^{-1}$).

compared to the linearly-varying α in Eq. 3.1. Additionally, the presence of clouds often contributes to a higher effective resolution and corresponding errors in AERIOE_FD compared to AERIOE. For the D05 methods, DESROZIERS_FD generally features a similar shape to DESROZIERS but with larger error magnitudes at some sites (e.g., moisture errors at FP4 in Fig. 5.7c).

The largest impact when utilizing the flow-dependent methods results from the increased temporal variance of the observation errors at all sites and for both temperature and moisture (see error bars in Fig. 5.6-5.7). AERIOE_FD features more variability than AERIOE throughout most of the profile due to the addition of γ in Eq. 5.1 that changes for each retrieval. AERIOE_FD also features much larger error variability near the top of the profile, likely due to the presence of clouds increasing the effective resolution of some retrievals. Next, given its relationship to the innovation and analysis residuals which can vary considerably for each DA cycle, DESROZIERS_FD produces the largest error variance. We note that no variation is shown for DESROZIERS since only a single, static error profile is computed for each site. The larger standard deviation of the flow-dependent methods (AERIOE_FD, DESROZIERS_FD) also more closely matches the observed variability in the RMSD between AERI retrievals and observed rawinsondes (Fig. 5.1; a proxy for the true observation error). This finding suggests that these advanced methods might better capture potential systematic errors wherein the observation uncertainty increases in specific meteorological conditions.

5.5 Impact of observation errors on the prediction of the 15 July 2015 MCS

Next, we evaluate the forecast impacts of assimilating AERI data using the above observation error profiles generated by each method. We primarily analyze the evolution of the 15 July MCS through NEP and FSS computed over a 16-km neighborhood. We

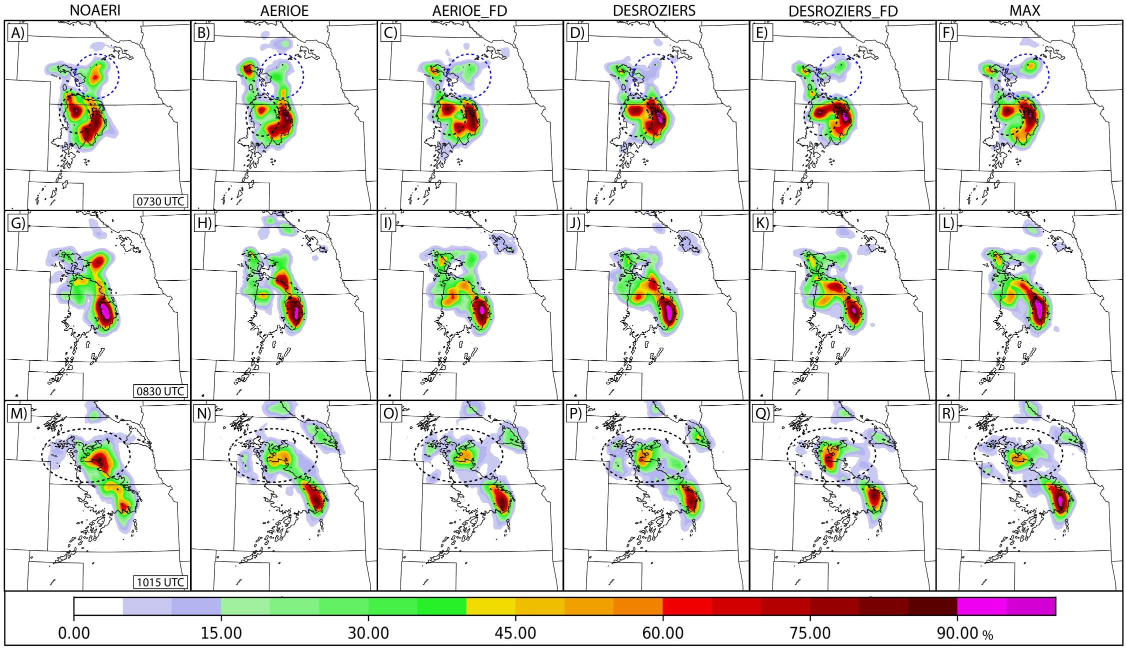


Figure 5.8: Neighborhood ensemble probabilities (NEP; %) for each observation error experiment and valid at (a-f) 0730 UTC; (g-l) 0830 UTC; and (m-r) 1015 UTC 15 July. The probabilities are computed for composite reflectivity exceeding 30 dBZ. Also overlaid in black are the 30 dBZ contours of observed composite reflectivity from MRMS. See text for a description of the circled regions.

note that the interpretation of these results does not change when verifying over different neighborhoods or forecast products (e.g., reflectivity or precipitation).

Without assimilating any AERI data, an MCS develops in the correct location with maximum probabilities located along the main convective line and in the trailing stratiform region in northwestern Kansas (Fig. 5.8a). Although NOAERI correctly predicts the early structure of the MCS, a large region of spurious precipitation develops to the north of the primary convective line (blue circle in Fig. 5.8a). This spurious convection, which we discuss at length throughout the rest of this chapter, eventually expands into a large area of precipitation and results in an apparent northern shift of the MCS relative to observations (circled region in Fig. 5.8m). Additionally, we note that NOAERI fails to maintain the MCS in Kansas such that ensemble probabilities along the observed convective line weaken to $<60\%$ by 1030 UTC (forecast hour 4; Fig. 5.8m).

Assimilating the AERI data using the static inflation method introduced in section 3.3.1 (AERIOE; Fig. 5.8b,h,n) slightly degrades the forecast quality during forecast hours 0-3. Primarily, the observation errors in AERIOE reduce the convective probabilities within the stratiform region of the MCS by $\sim$30% (southern black circles in Fig. 5.8a,b). AERIOE also displaces, but does not suppress, much of the spurious convection further south such that it connects with the main convective line of the MCS (Fig. 5.8b). This shift causes the MCS to orient increasingly north-south and again extend too far north in AERIOE. By 1015 UTC (Fig. 5.8m,n), AERIOE eventually begins to indicate a better forecast than NOAERI due to improvements in both the spurious precipitation (reduced probabilities by $\sim$20%) and better maintenance of the convective line in the MCS (increased probabilities by $\sim$20%). However, the early forecast degradations within the trailing stratiform region, along with little improvement in the spurious precipitation, contribute to AERIOE producing a lower mean FSS than NOAERI for the weaker precipitation threshold (Table 5.3).

While the static inflation method slightly degrades the forecast relative to the baseline simulation, its flow-dependent extension (AERIOE_FD; Fig. 5.8c,i,o) improves upon many of the above-mentioned problems. Primarily, we find that AERIOE_FD results in reduced probabilities for the spurious precipitation across most of southern Nebraska (Fig. 5.8c,i). AERIOE_FD also better maintains the stratiform region of the MCS as indicated by higher probabilities in northwestern Kansas from 0730-0830 UTC (Fig. 5.8c,i). These results contribute to an overall more skillful forecast in AERIOE_FD compared to NOAERI and AERIOE (Table 5.3) and suggest that some of the forecast degradations discussed in Chapter 4 can be alleviated by assigning flow-dependent observation errors for remote sensing data. However, the forecast in AERIOE_FD eventually converges with AERIOE after 1030 UTC (not shown), indicating that the impact from modifying the observation errors only lasts $\sim$4 h.

This study also represents one of the first known works to evaluate the impact of using the D05 method to assign observation errors for convective-scale DA and prediction. The static implementation of this method, DESROZIERS (Fig. 5.8d,j,p), similarly improves upon many of the issues discussed above. DESROZIERS simulates less spurious precipitation in central Nebraska and enhances probabilities within the trailing stratiform region of the MCS compared to both AERIOE and AERIOE_FD. However, DESROZIERS still develops low probabilities for unobserved convection along the Nebraska-Kansas border that leads to a slight misorientation of the MCS at later lead times (Fig. 5.8j). While the spurious precipitation is better suppressed in AERIOE_FD, DESROZIERS instead shows larger improvements to the convective line of the MCS as indicated by higher probabilities in north-central Kansas from 0730-0830 UTC (Fig. 5.8d,j).

Finally, the flow-dependent extension of the D05 method (DESROZIERS_FD; Fig. 5.8e,k,q) produces large improvements upon DESROZIERS (Table 5.3) and produces the best forecast of any experiment apart from MAX. At 0730 UTC, DESROZIERS_FD accurately simulates nearly all the convection in Nebraska and only fails to capture a weak region of precipitation in northeast Nebraska that no experiment simulates until later. DESROZIERS_FD also produces the best forecast for the MCS in Kansas, including high probabilities both within the convective line and in the stratiform region (Fig. 5.8e,k). At 0830 UTC, the bowing structure of the MCS is well captured in DESROZIERS_FD as indicated by the northwest-southeast orientation in convective probabilities just north of the Kansas-Nebraska border (Fig. 5.8k).

Table 5.3 summarizes this NEP analysis and reveals three major results, including: (1) each method, apart from AERIOE, results in a positive impact for the convective forecast when assimilating AERI data compared to NOAERI; (2) the experiments based on the observation-space diagnostics in D05 (DESROZIERS, DESROZIERS_FD) produce higher skill than the inflation of the AERIoe retrieval errors (AERIOE, AERIOE_FD); and (3) the flow-dependent extensions of each method produce higher skill than their static counterparts. Each result is primarily related to improved suppression of spurious convection, with a small contribution stemming from increased probabilities within the MCS. Though not discussed

Table 5.3: Fractions Skill Score (FSS) values computed for accumulated precipitation exceeding 2.54 $mm\ hr^{-1}$ and 6.35 $mm\ hr^{-1}$. Each score is computed across the entire forecast domain (d02 in Fig. 5.5) using a 16-km neighborhood and then averaged over the 7-h forecast period. The verifying precipitation observations are obtained from MRMS.

Experiment	FSS (2.54 $mm\ hr^{-1}$)	FSS (6.35 $mm\ hr^{-1}$)
NOAERI	0.49	0.54
AERIOE	0.47	0.56
AERIOE_FD	0.50	0.61
DESROZIERS	0.50	0.60
DESROZIERS_FD	0.51	0.61
MAX	0.51	0.63

above, we also find that MAX (Fig. 5.8f,l,r) produces a subjectively similar forecast to DESROZIERS_FD but with higher probabilities within the main convective line (Fig. 5.8l). As such, MAX produces the most skillful forecast compared to other experiments (Table 5.3). This finding suggests that both sets of methods have individual benefits that are optimized when utilized together.

5.5.1 Analysis of spurious precipitation

Given that modifications to the spurious precipitation in Nebraska are responsible for much of the skill differences between the experiments, we perform an additional analysis to determine what mechanisms lead to the spurious CI. We compare member 1 from AERIOE and MAX for this analysis, as these two simulations produce the most and least spurious convection between the experiments that assimilate AERI data, respectively (Fig. 5.9). Member 1 also best illustrates the differences highlighted in the previous section, as AERIOE generates a north-south oriented line of unobserved convection (Fig. 5.9a,c) whereas MAX produces a small amount of east-west oriented convection that is observed at earlier lead times (Fig. 5.9b,d). As with DESROZIERS_FD, MAX instead produces a secondary northwest-southeast oriented convective event that matches the observed system well (circled in Fig. 5.9f). This secondary CI event in MAX occurs due to northerly winds converging with outflow associated with the MCS (yellow circle in Fig. 5.10b). Due to the large amount of spurious convection inhibiting the northwesterly winds in southwestern Nebraska (Fig. 5.10a), AERIOE does not simulate the convergence and associated convective band. As such, we hypothesize that the improved structure of the MCS in DESROZIERS_FD and MAX also results from improved suppression of the spurious convection.

We find that the spurious convection develops within a north-south oriented midlevel moist layer located along the Kansas-Nebraska border (Fig. 5.11a). This moisture bulge results in enhanced midlevel instability, as denoted by air parcels at 1 km AGL featuring $\sim$3000 $J\ kg^{-1}$ of CAPE (Fig. 5.11c). In addition to the enhanced instability, this region also features low-level convergence due to the strong westerly winds associated with the MCS outflow (Fig. 5.11a). The spurious convection of interest is primarily supported by this enhanced instability and convergence. Although the convergence ahead of the MCS is predicted similarly in MAX, using the flow-dependent observation errors produces a weaker moist layer that is displaced further south (Fig. 5.11b). We note that a similar moisture bulge is seen in RAP analyses at 0500 UTC (Fig. 5.2a) but does not extend as far north as is simulated in AERIOE (Fig. 5.11a), further supporting the corrections shown in MAX. As such, MAX simulates lower MUCAPE in southern Nebraska, leading to a suppression of the spurious convection relative to AERIOE (Fig. 5.11d).

5.6 Impact of observation errors during DA cycling

In this final section of results, we relate improvements in the convective forecast to individual observation error profiles. We present difference and ensemble correlation plots to evaluate the impacts specifically for the spurious convection and the trailing stratiform region. Note that the differences in Fig. 5.12, apart from the first column, are plotted relative to AERIOE as that method has been applied for our previous studies (Chapters 3, 4) and can be treated as a baseline assimilation method.

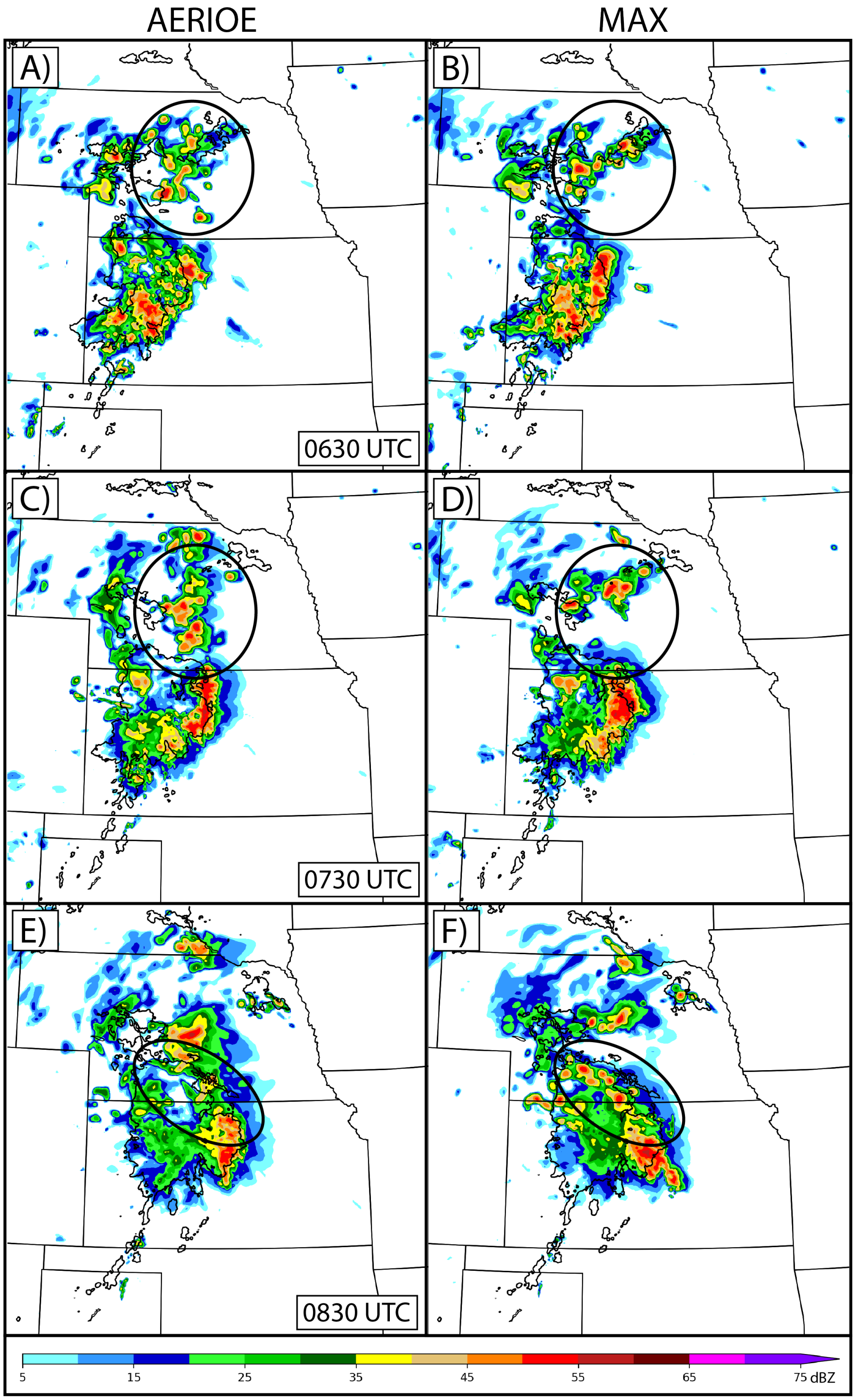


Figure 5.9: Simulated composite reflectivity from member 1 of AERIOE and MAX valid at (a,b) 0630 UTC; (c,d) 0730 UTC; and (e,f) 0830 UTC. Also overlaid in black are the 30 dBZ contours of observed composite reflectivity from MRMS. See text for a description of the circled regions.

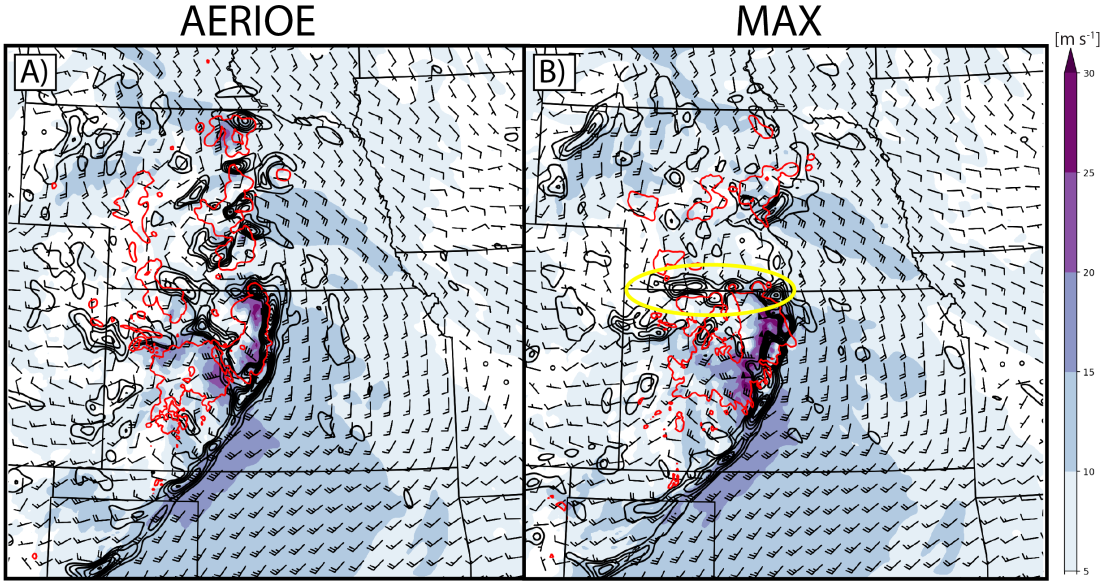


Figure 5.10: Simulated 850 hPa winds (ms^{-1}; shading and barbs), horizontal convergence (contoured in black every $+5^{-6}$ s^{-1}), and composite reflectivity (30 dBZ contours in red) from member 1 of AERIOE and MAX. Each plot is valid at 0730 UTC 15 July. The half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. The yellow circle in (b) indicates convergence band that results from northerly winds converging with outflow.

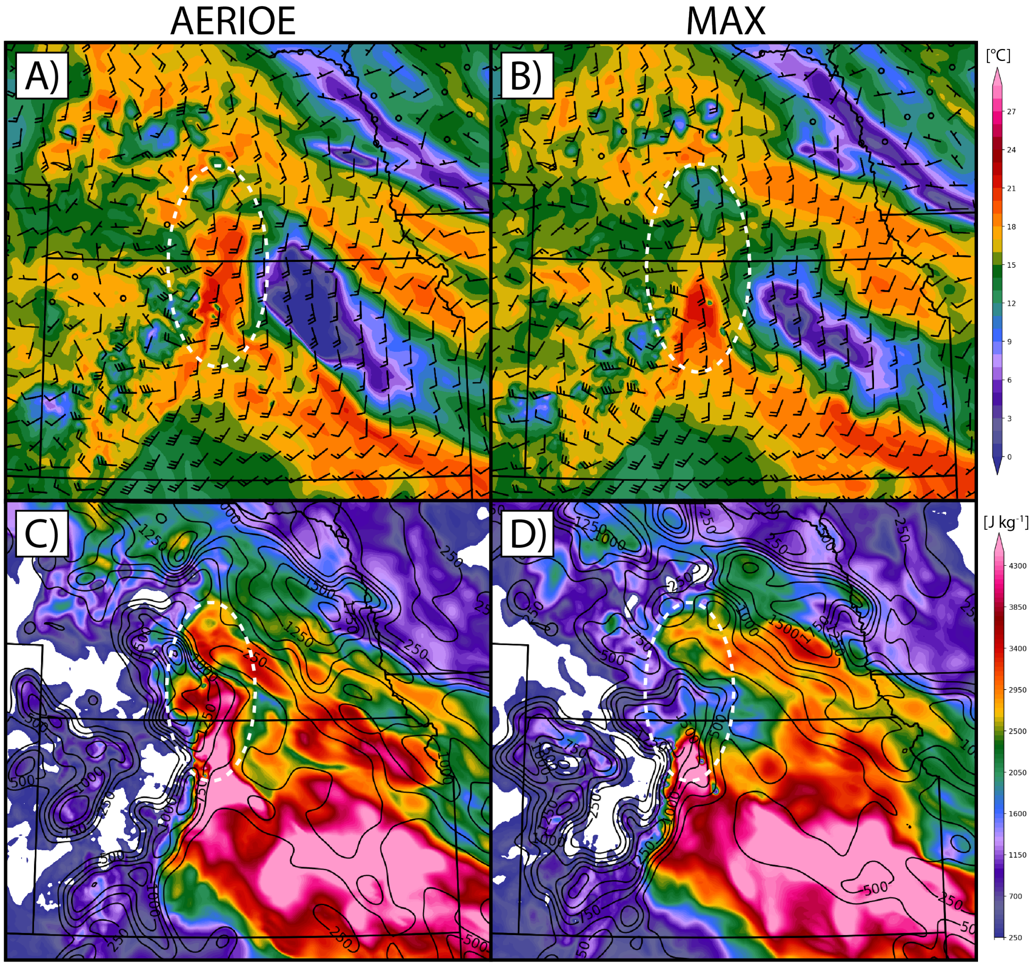


Figure 5.11: Simulated pre-convective thermodynamic fields from member 1 of the AERIOE and MAX experiments, including (a,b) 850 hPa dewpoint temperatures (°C; shading) and winds (ms^{-1}; barbs); and (c,d) most unstable CAPE ($J\ kg^{-1}$, shading) and lifted parcel levels for the most unstable parcels (m AGL; black contours). In (a,b), the half barbs represent wind speeds of 2.5 ms^{-1} and the full barbs represent wind speeds of 5 ms^{-1}. The white circles represent regions corresponding to the spurious convection discussed in the text.

5.6.1 Impact to spurious precipitation

The difference plots reveal two moisture surges during DA that contribute to the spurious convection simulated in AERIOE. The first of these surges occurs at 0100 UTC just east of FP3 when AERIOE increases the 850 hPa mixing ratio by $\sim 3\ g\ kg^{-1}$ relative to NOAERI (red circle in Fig. 5.12a). This region of enhanced moisture is advected to the northwest between 0100-0530 UTC. At 0400 UTC, a second region of enhanced moisture develops in AERIOE, again just east of FP3 (Fig. 5.12f). These two moisture surges merge during later DA cycles, eventually producing a convective-scale region of enhanced moisture compared to NOAERI (+1.03 $g\ kg^{-1}$, red box in Fig. 5.12k). The final location of this positive moisture increment corresponds to the enhanced MUCAPE shown in Fig. 5.11a,c.

Assimilating the retrievals in experiments other than AERIOE increasingly reduces the impact of both moisture surges (see negative values in the red boxes of Fig. 5.12l-o). The moistening is least reduced in AERIOE_FD (-1.52 $g\ kg^{-1}$) and DESROZIERS (-2.21 $g\ kg^{-1}$), and most reduced in DESROZIERS_FD (-2.92 $g\ kg^{-1}$) and MAX (-2.80 $g\ kg^{-1}$). We note that in general, each difference field features the same spatial structure as (AERIOE – NOAERI) except for the opposite sign, indicating that the various methods only modify the magnitude of the observation impacts. We also summarize the moisture impacts corresponding to the spurious convection by computing average soundings during the final DA cycle (Fig. 5.13a). Here, each experiment reduces MUCAPE and enhances MUCIN compared to AERIOE. Again, AERIOE_FD features the smallest relative impact while DESROZIERS_FD and MAX produce the largest impact. Given the modifications to the convective indices, these relative differences likely explain why AERIOE_FD and DESROZIERS reduce the amount of spurious convection while DESROZIERS_FD and MAX almost entirely suppress it.

Next, we perform an ensemble correlation analysis to determine which AERI retrievals contribute most to the positive moisture increments. While both moisture surges occur near FP3, Fig. 5.14 reveals that the impacts are likely related to AERI observations assimilated

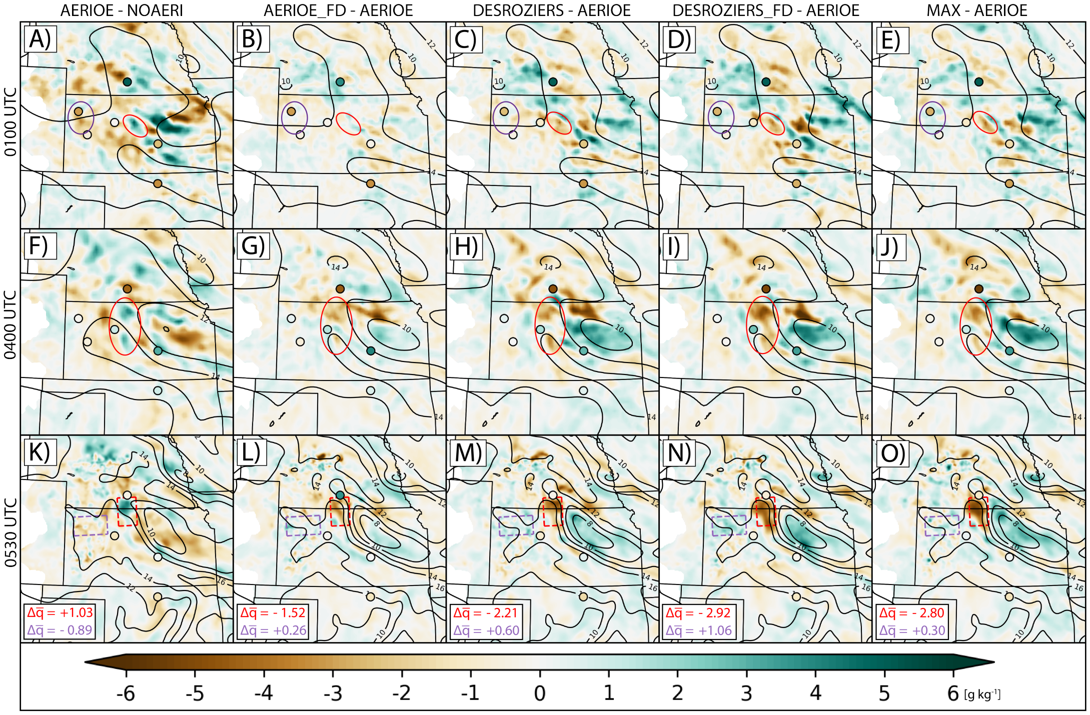


Figure 5.12: Ensemble mean differences of 850 hPa water vapor mixing ratio ($g\ kg^{-1}$) for (a,f,k) AERIOE minus NOAERI; (b,g,l) AERIOE_FD minus AERIOE; (c,h,m) DESROZIERS minus AERIOE; (d,I,n) DESROZIERS_FD minus DESROZIERS; and (e,j,o) MAX minus AERIOE. The black contours indicate the ensemble mean 850 hPa water vapor mixing ratio simulated by AERIOE. Also overlaid are the innovations for AERI moisture retrievals closest to 850 hPa and assimilated at each cycle using the same color scale. The red regions correspond to positive moisture increments in AERIOE that enhance spurious convection in southern Nebraska, and the purple regions correspond to negative moisture increments in AERIOE that degrade the trailing stratiform region of the 15 July MCS. In (k-o), a mean difference value corresponding to the dashed rectangles are annotated.

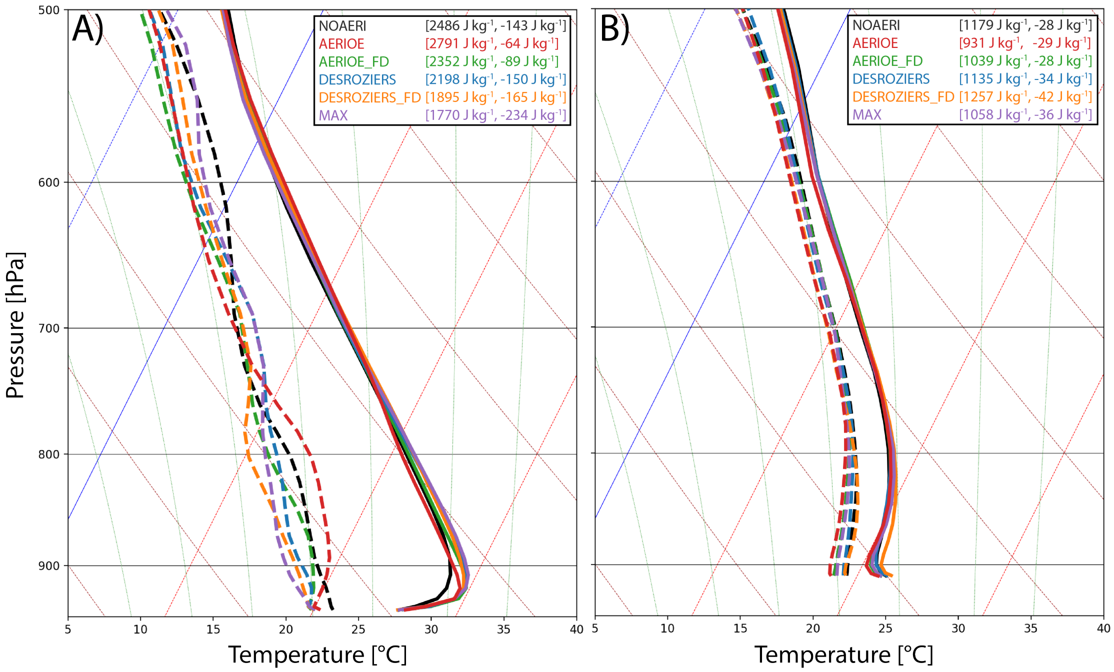


Figure 5.13: Ensemble mean soundings computed over the (a) dashed red box; and (b) dashed purple box in Fig. 5.12k-o at 0530 UTC 15 July. Also annotated on each sounding are the most unstable CAPE ($J\ kg^{-1}$; first metric in brackets) and CIN associated with the most unstable parcel ($J\ kg^{-1}$; second metric in brackets).

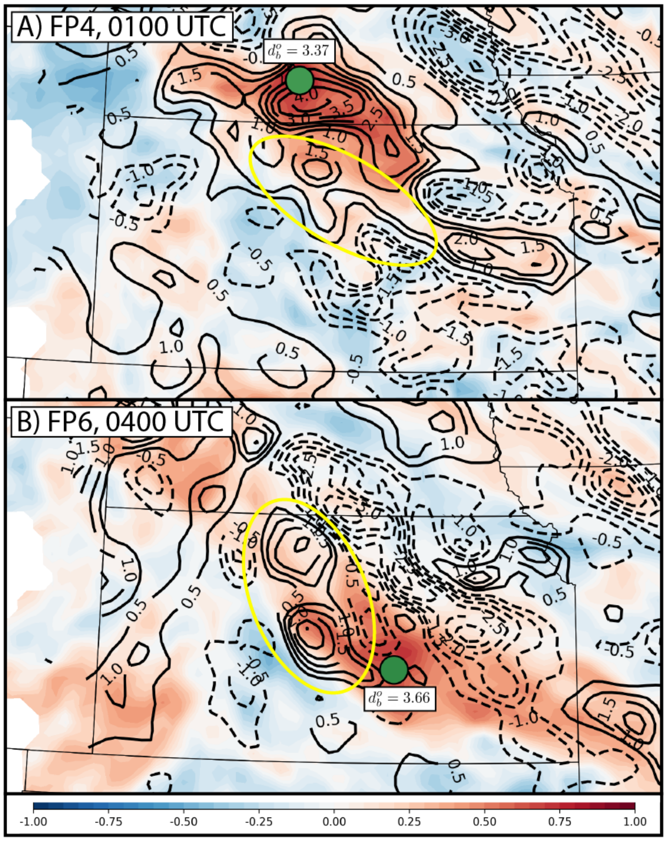


Figure 5.14: Background ensemble correlations between 850 hPa water vapor mixing ratio at an AERI platform and the rest of the domain. The plots are computed at (a) FP4 valid at 0100 UTC; and (b) FP6 valid at 0400 UTC. The location of the (a) FP4; and (b) FP6 sites are annotated with the green dots, and the mean values of their 850 hPa moisture innovations are overlaid (d_b^o; $g\ kg^{-1}$). Also overlaid in black contours are the analysis increments for the AERIOE experiment. See text for a description of the yellow circles.

from FP4 (Fig. 5.14a; 0100 UTC surge) and FP6 (Fig. 5.14b; 0400 UTC surge). Both sites feature very moist midlevels (innovation $> 3\ g\ kg^{-1}$), meaning that assimilating these data results in strong, positive increments in regions of positive correlation and strong, negative increments in regions of negative correlation. Though the analysis increments are a superposition of all sites and observation levels assimilated, the correlation structures at FP4 and FP6 roughly correspond to the moisture surges highlighted above (yellow circles in Fig. 5.14).

Finally, we present the AERI retrievals assimilated at FP4 and FP6 to better understand how the observation errors relate to these moisture impacts (Fig. 5.15-5.16). We note that the background and analysis profiles are only shown for AERIOE as the corresponding profiles for other experiments do not significantly differ. As with Fig. 5.14, both retrievals indicate positive innovations due to the very moist retrievals above 900 hPa relative to the background. AERIOE indicates the lowest observation error at both sites (Fig. 5.15c, Fig. 5.16c) and thus assigns the largest weight to the moist retrievals during DA. AERIOE_FD features only slightly larger errors than AERIOE due to a marginally larger vertical resolution compared to the median (Fig. 5.15b, Fig. 5.16b). Despite this, the small error increases in AERIOE_FD appear to have been significant enough to reduce the impact of the moistening. Conversely, both DESROZIERS and DESROZIERS_FD produce much larger observation errors around 850 hPa, with DESROZIERS_FD reaching 4.5 $g\ kg^{-1}$ at FP6 (Fig. 5.16c) due to the very large innovation and analysis residuals (Fig. 5.16a). These larger errors reduce the weight assigned to the moist AERI retrievals and thus support the significant reduction in spurious convection in DESORIZERS and DESROZIERS_FD. Apart from the improved forecast results when assimilating the larger observation errors, the increased errors are further supported by a collocated rawinsonde launched at FP6 that indicates a strong moisture bias in the AERI retrievals (Fig. 5.16a). As the observation errors for MAX are approximately equal to DESROZIERS_FD for the FP4 and FP6 retrievals, these error profiles also demonstrate why MAX similarly suppresses the spurious convection in central Nebraska as in DESROZIERS_FD.

5.6.2 Impact to trailing stratiform precipitation

Finally, we also analyze the impacts of the observation errors on the low-level moisture near the trailing stratiform region of the 15 July MCS (purple regions in Fig. 5.12k-o). Although the small NEP increases in this region do not impact the skill as much as the improvements to the spurious precipitation, each method still shows a measurable benefit compared to the original static inflation method applied in Chapters 3-4.

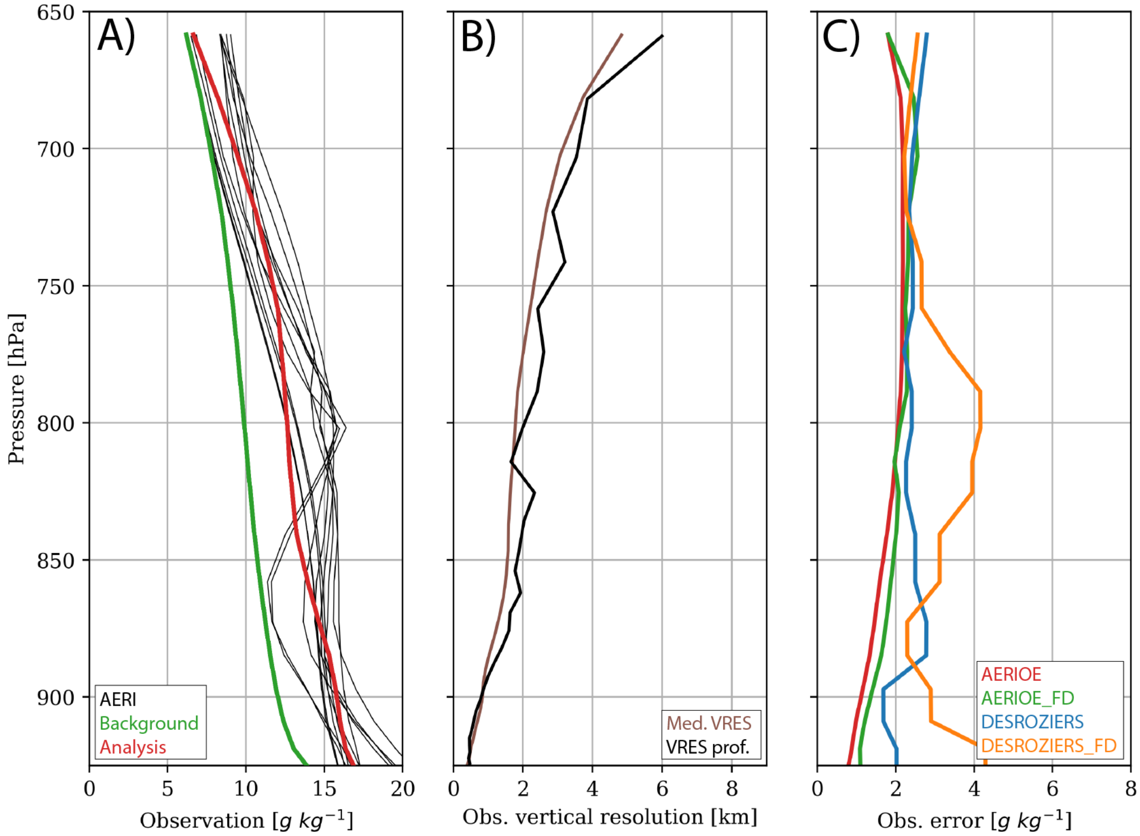


Figure 5.15: (a) AERI water vapor mixing ratio ($g\ kg^{-1}$) observations assimilated at FP4 during the 0100 UTC cycle, along with corresponding background and analysis profiles from the AERIOE experiment. (b) Median effective vertical resolution (km) profiles output by the AERIoe retrieval at the same time. The black profile indicates the median computed over all PECAN retrievals, while the brown profile indicates the median computed only over the profiles assimilated during this cycle. (c) Observation error ($g\ kg^{-1}$) profiles for these data assigned in each experiment.

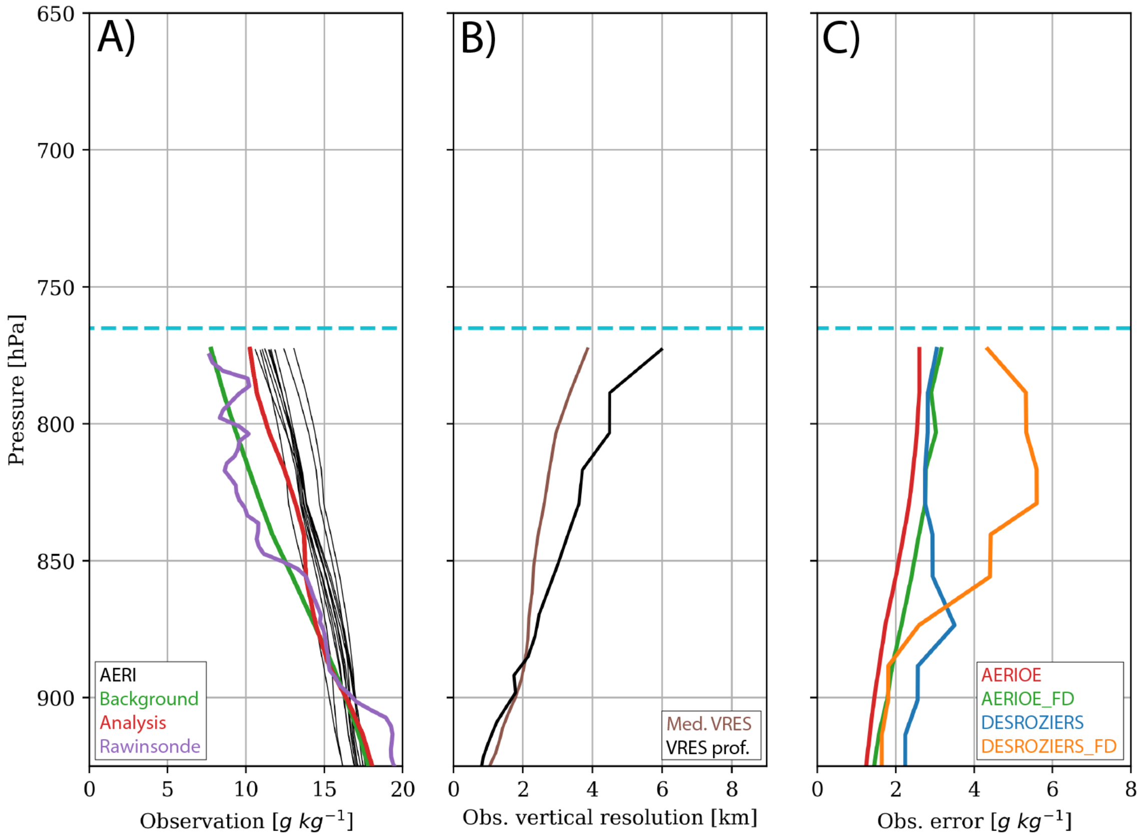


Figure 5.16: As in Fig. 5.15 but for the FP6 moisture observations assimilated at 0400 UTC. Also overlaid now are the cloud base height indicated by a collocated lidar or ceilometer (dashed cyan line), and a corresponding moisture profile collected by a collocated rawinsonde (purple line in [a]).

In general, we find that assimilating the AERI retrievals in the baseline AERIOE experiment produces moderate drying in northwestern Kansas, especially during the 0100 UTC cycle near FP5 (Fig. 5.12a) and the 0415 UTC cycle near SPARC (not shown). These moisture impacts are mostly stationary and eventually lead to a moderate drying effect underneath the ongoing MCS during the final DA cycle (-0.89 $g\ kg^{-1}$, Fig. 5.13k). The moisture impacts are also illustrated by computing an average sounding that reveals reduced dewpoint temperatures below 600 hPa and 200 $J\ kg^{-1}$ less of MUCAPE in AERIOE compared to NOAERI (Fig. 5.13b). The reduced instability likely explains why convection in the trailing stratiform region decays in AERIOE compared to other experiments. Conversely, assigning observation errors using the other three methods reduces the impact of this drying (Fig. 5.12l-o) and correspondingly increases the MUCAPE (Fig. 5.13b). As before, the relative magnitude of these differences is smallest in AERIOE_FD (Fig. 5.12l), moderate in DESROZIERS (Fig. 5.12m), and largest in DESROZIERS_FD (Fig. 5.12n). Unlike the previous analysis, we find here that MAX produces a smaller impact for the moisture in the stratiform region compared to DESROZIERS_FD (Fig. 5.12o).

We do not present a correlation analysis for this section, as the moisture impacts near the trailing stratiform region occur directly above FP5 and SPARC (Fig. 5.17-5.18). Unlike the previous section which details moist biases, these AERI observations instead feature dry layers that are not supported by collocated rawinsondes. Specifically, the 0100 UTC retrieval at FP5 is too dry between 875-750 hPa (Fig. 5.17a), while the 0415 UTC retrieval at SPARC is too dry above 850 hPa (Fig. 5.18a). At FP5, AERIOE and AERIOE_FD generally feature similar observation errors and thus assign the heaviest weight to these dry biases (Fig. 5.17c). Due to the large residuals at FP5 between 850-750 hPa, DESROZIERS_FD again diagnoses the largest observation errors to reduce the impact of the drying (Fig. 5.17c). This midlevel error increase is also partially represented by the static DESROZIERS experiment, indicating a potential bias in the moisture retrievals from FP5. At SPARC, the dry bias occurs just below cloud base, indicating that it could potentially be related

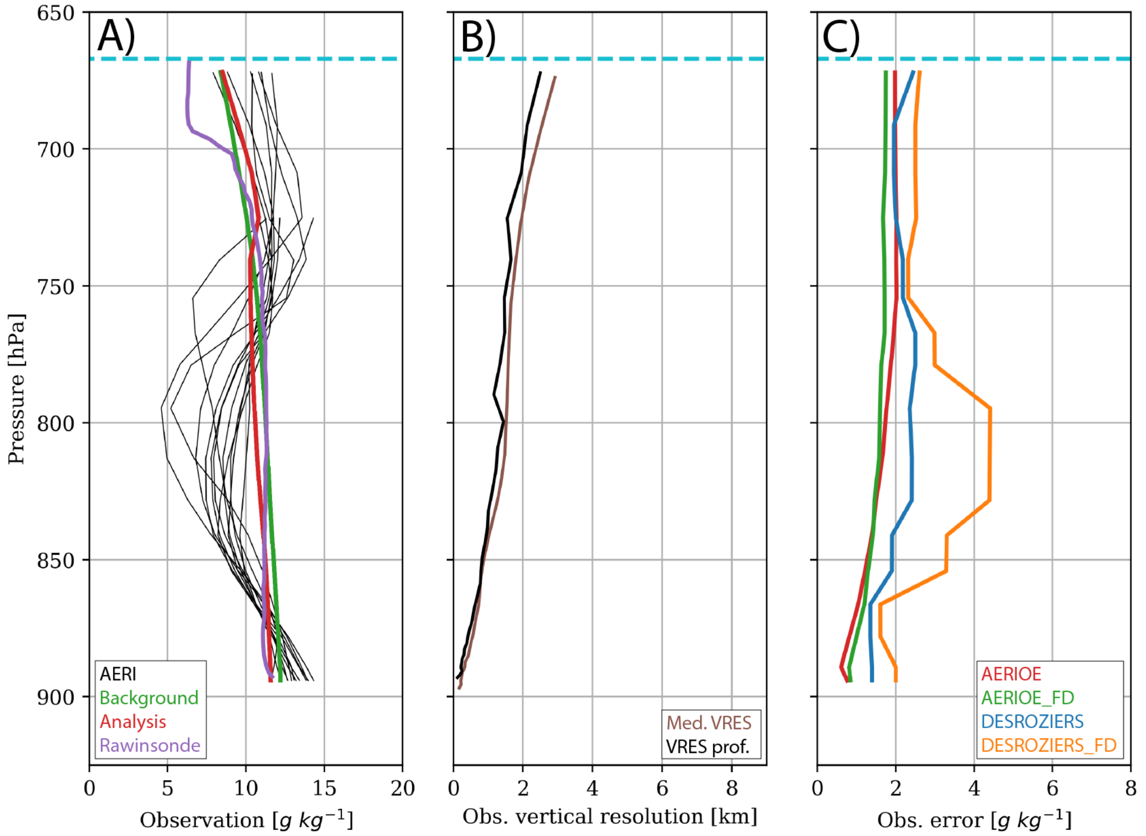


Figure 5.17: As in Fig. 5.16 but for the FP5 moisture observations assimilated at 0100 UTC.

to the emission issues detailed in Turner and Blumberg (2019). The issue is represented in the vertical resolution profile above 850 hPa (Fig. 5.18b) and as such, the observation errors are correspondingly inflated in AERIOE_FD (Fig. 5.18c). Conversely, DESROZIERS and DESROZIERS_FD produce much lower observation errors here due to the near-zero analysis residuals throughout much of the profile (Fig. 5.18a,c).

5.7 Discussion

This dissertation chapter evaluates the impact of using static and flow-dependent methods for assigning observation errors in a convective-scale DA system. Using high-frequency thermodynamic profiles collected by AERIs, we compare two sets of experiments including a method that inflates retrieval errors, and another that computes the full observation error

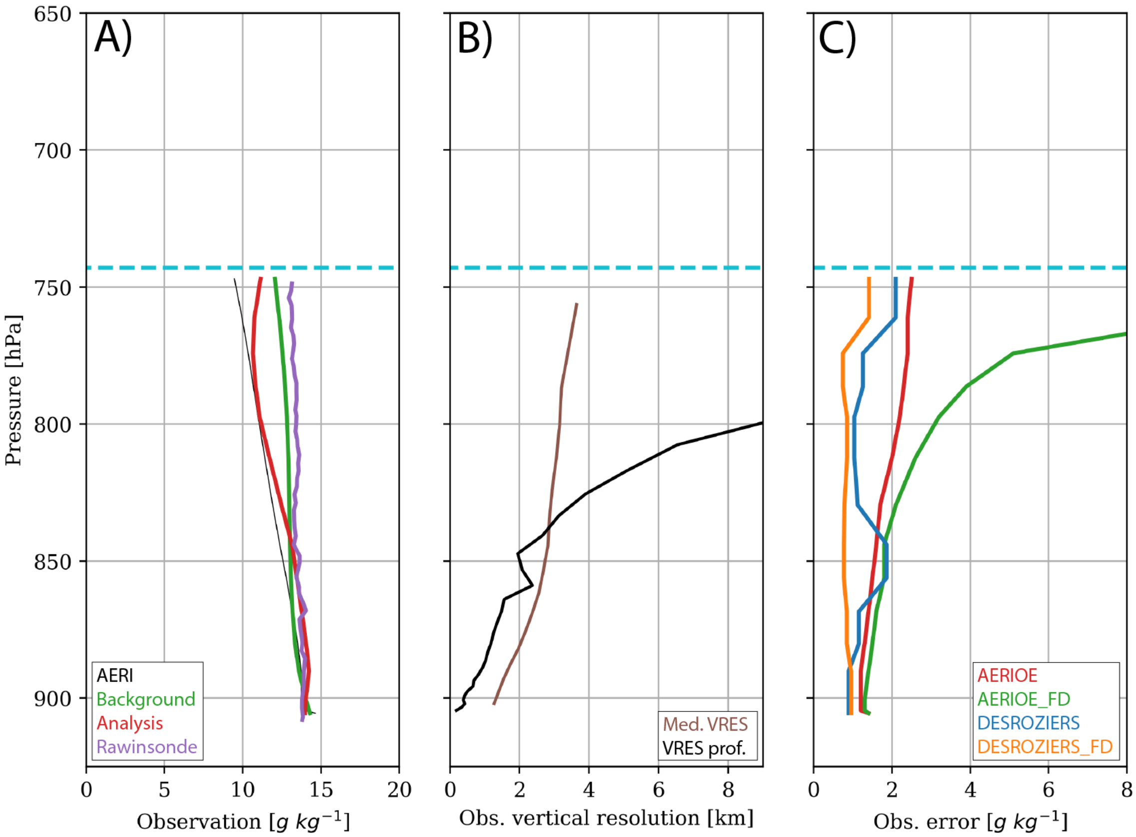


Figure 5.18: As in Fig. 5.16 but for the SPARC moisture observations assimilated at 0415 UTC.

based on the D05 diagnostic. In general, the inflation-based methods produce a smooth error profile that increases with height, while the D05 methods produce larger errors in the boundary layer. When the static inflation and D05 methods are extended to be flow-dependent, the latter produce more variable errors that better match the wide distribution of observed statistics. Additionally, the flow-dependent application of the inflation method produces larger errors below cloud layers likely due to issues separating cloudy emissions from atmospheric emissions (Turner and Blumberg 2019).

When assimilating the AERI data using these observation errors, we find that the static error inflation applied in previous studies can reduce skill compared to a reference experiment that does not assimilate any profiling data. Specifically, the static error inflation results in reduced probabilities within the trailing stratiform region of the 15 July MCS due to high weights assigned to retrievals with a dry bias. However, these forecast degradations are removed when assimilating AERI data using the flow-dependent inflation method. This extension improves the probabilities within the trailing stratiform region and suppresses spurious convection developing near the MCS. These results indicate that constant observation errors or static inflation methods are likely inaccurate for remote sensing datasets and that careful attention needs to be paid to how these errors are assigned.

We find additional forecast improvements when assimilating thermodynamic data using the D05 method relative to the inflation method. While the static implementation of D05 demonstrates advantages over the inflation methods, a flow-dependent implementation of this method provides further improvements due to its ability to increase observation errors for any retrieval that differs strongly from the background or analysis ensemble. However, the advantages of the flow-dependent implementation over the static version of D05 are small, suggesting that such benefits might not be worth the increased cost of computing multiple analyses per DA update. Furthermore, the similarity between the two implementations suggest that AERI retrievals feature somewhat constant systematic errors that could perhaps be better controlled through a bias correction scheme.

In general, we find that most of these benefits result from increasing the observation errors relative to the baseline (i.e., decreasing weights) for retrievals with a moisture bias. This result is exemplified by MAX producing the highest skill compared to any experiment and indicates that larger errors are often required to account for potential systematic observation errors. Conversely, we instead find little evidence of improvements from adaptively assigning a lower observation error. Additionally, we find that the convective impacts are largely related to changes to the moisture error profile as opposed to temperature errors. This follows Peters et al. (2017) who document the strong sensitivity of MCS forecast errors to moisture errors. We also note larger impacts to the moisture field when assimilating thermodynamic profilers in Chapters 4 and 5.

This study constitutes a proof-of-concept work for using flow-dependent methods and the D05 diagnostic for assigning observation errors for convective-scale DA. However, by definition, the flow-dependent component of observation errors (representation and systematic errors) varies considerably by case and observation type. As such, we plan to continue this work using additional observation types including Doppler lidars and satellite data, both of which are also known to feature flow-dependent observation errors (e.g., Chapter 4; Fielding and Stiller 2019). Finally, a large advantage of the D05 method is its ability to diagnose a full covariance matrix including observation error correlations. Though horizontal error correlations are likely small for profiling instruments (i.e., site to site correlation), vertical correlations can be large (Turner and Löhnert 2014). Thus, our future work will also include adopting other DA methods such as EnVar (e.g., Wang et al. 2013; Wang and Lei 2014) to evaluate the impact of assimilating the full $\widetilde{\mathbf{R}}$ diagnosed by the D05 method instead of only inflating the observation error variances.

Chapter 6

Conclusions

Convection-allowing NWP systems typically perform poorly when predicting elevated, nocturnal thunderstorms compared to surface-based convection. This deficiency partially results from the inability of current models to correctly simulate the mechanisms responsible for generating and sustaining nocturnal convection (e.g., the nocturnal LLJ, elevated moist layer, atmospheric bores, etc.). Additionally, initial condition errors contribute to the poor forecast skill of NWP when predicting nocturnal CI and the evolution of MCSs at night. These initial condition errors result from the lack of high-frequency profiling observations that sample important elevated mechanisms. To improve the prediction and understanding of nocturnal convection, scientists conducted the PECAN field campaign during 2015 (Geerts et al. 2017). As part of PECAN, a large network of fixed and mobile instruments was deployed that included ground-based remote sensing profilers and high-frequency rawinsonde and surface observations. This dissertation conducts detailed experiments to explore whether assimilating these novel datasets can improve predictions of nocturnal convection.

We first evaluate the impact of assimilating the PECAN dataset for a single nocturnal CI event observed during PECAN on 26 June 2015. Compared to operational forecasts, assimilating the PECAN dataset improves the timing, location, and orientation of the CI event. We also find that the largest forecast improvements result when assimilating these data on a 12-km grid as opposed to the inner, convection-allowing domain. These results suggest that mesoscale errors heavily contribute to reduced forecast skill and that even a coarse network of profilers can alleviate such problems. By conducting experiments in the data denial framework, we find that assimilating AERIs, RWPs, and high-frequency rawinsondes produces the largest and most sustained impact by enhancing the elevated moist layer in the region of CI. Assimilating Doppler lidar and surface observations also improve the CI forecast, though the impacts are not sustained for as long as other observation types.

Motivated by the diverse forcing mechanisms and environments that can generate nocturnal convection, we next expand our work into a systematic evaluation of the impact of assimilating remote sensing profilers for forecasts of nocturnal CI. Given that Stalker et al. (2013) propose implementing new observation networks alongside pre-existing infrastructure, we also modify our methods such that we only evaluate the impact of assimilating collocated datasets.

Our systematic study shows that assimilating remote sensing, thermodynamic profilers again produces small but consistent improvements across a larger set of cases. Assimilating these data contributes to improved forecasts by enhancing the elevated moist layer prior to CI. These results confirm a systematic benefit for nocturnal CI forecasts when assimilating ground-based thermodynamic profilers. Given the relatively small number of instruments assimilated here, it is likely that expanding the network could lead to even further improvements. Conversely, assimilating the kinematic data often degrades the forecast performance for the detection of nocturnal CI, likely due to opposite impacts to the pre-convective thermodynamic fields compared to other experiments. Despite this, assimilating both datasets together leads to slight improvements to various CI properties (e.g., timing, location, and orientation).

We hypothesize that some of the forecast degradations from assimilating remote sensing data could be a result of underestimated observation errors at some sites. This problem occurs when using a static method to inflate the provided instrument error and thus assigning very similar errors across the entire network. Such methods are not likely accurate given that remote sensing datasets are known to feature strong flow-dependent components to their observation errors (Fielding and Stiller 2019). To further test this hypothesis, we design our final set of experiments that evaluates the impact of utilizing novel methods to assign observation errors for remote sensing datasets. We compare static to flow-dependent methods, along with methods that inflate the retrieval errors against methods that compute the full observation error variance.

Evaluating these methods for a nocturnal MCS on 15 July 2015 confirms that using the static error inflation method can degrade convective forecasts relative to an experiment where no remote sensing data are assimilated. We find that adaptively inflating the observation errors instead can alleviate some problems and improve the convective forecast. Further improvements occur when using the Desroziers et al. (2005) method that computes the full observation errors based on observation-space diagnostics. A flow-dependent extension of the Desroziers et al. (2005) methods leads to some additional improvements, though the limited benefits might not be worth the increased cost. All of the above benefits are primarily linked to adaptively increasing the observation errors for problematic retrievals. This proof-of-concept study suggests a large sensitivity of observation impacts to the assigned observation errors and motivates future work in this area for other remote sensing instruments (e.g., satellite-based profilers).

This dissertation is partially motivated by the Nationwide Network of Networks vision (National Research Council 2009) that aims to identify potential observation networks to improve forecasts of high-impact weather impacts. We show here that ground-based thermodynamic profilers are especially impactful and result in the largest improvements to the forecast skill for nocturnal convection compared to other instruments. Like previous works (e.g., Coniglio et al. 2019; Chipilski et al. 2020; Hu et al. 2019), we find that the main advantage of the thermodynamic profilers is their ability to detect frequent changes in elevated moisture and instability. The impacts of these instruments can be further improved through the proper assignment of their observation errors. While some aspects of their assimilation remain unaddressed (e.g., diagnosis of correlated observation errors), these results presented herein show that a nationwide network of AERIs could greatly improve the capabilities of convective-scale NWP systems to predict severe convection.

Bibliography

Adams-Selin, R., and R. S. Schumacher, 2019: Observations of low-frequency gravity waves during the PECAN field campaign. *18th Conf. on Mesoscale Meteorology*, Savannah, GA, Amer. Meteor. Soc., 3.5.

Banta, R. M., R. K. Newsom, J. K. Lundquisi, Y. L. Pichugina, R. L. Coulter, and L. Mahrt, 2002: Nocturnal low-level jet characteristics over Kansas during Cases-99. *Bound.-Layer Meteor.*, **105**, 221–252.

Benjamin, S. G., B. D. Jamison, W. R. Moninger, S. R. Sahm, B. E. Schwartz, and T. W. Schlatter, 2010: Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and mesonet observations via the RUC hourly assimilation cycle. *Mon. Wea. Rev.*, **138**, 1319–1343.

Benjamin, S. G., B. E. Schwartz, E. J. Szoke, and S. E. Koch, 2004: The value of wind profiler data in U. S. weather forecasting. *Bull. Amer. Meteor. Soc.*, **85**, 1871–1886.

Benjamin, S. G., and Coauthors, 2016: A north american hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Wea. Rev.*, **144**, 1669–1694.

Berger, 2004: Satellite wind superobbing. NWPSAF–MO–VS–016, 33 pp, https://www.ssec.wisc.edu/ howardb/Papers/superob_nwpsaf_final.pdf.

Blumberg, W. G., T. J. Wagner, D. D. Turner, and J. Correia, 2017: Quantifying the accuracy and uncertainty of diurnal thermodynamic profiles and convection indices derived from the Atmospheric Emitted Radiance Interferometer. *J. Appl. Meteor. Climatol.*, **56**, 2747–2766.

Bonner, W. D., 1968: Climatology of the low level jet. *Mon. Wea. Rev.*, **96**, 833–850.

Bormann, N., M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally, 2016: Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Quart. J. Roy. Meteor. Soc.*, **142**, 1767–1780.

Bormann, N. A., A. J. Geer, and P. Bauer, 2011: Estimates of observation-error characteristics in clear and cloudy regions for microwave imager radiances from numerical prediction models. *Quart. J. Roy. Meteor. Soc.*, **137**, 2014–2023.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part i: Description and single-observation experiments. *Mon. Wea. Rev.*, **138**, 1550–1566.

Burghardt, B. J., C. Evans, and P. J. Roebber, 2014: Assessing the predictability of convection initiation in the high plains using an object-based approach. *Wea. Forecasting*, **29**, 403–418.

Burlingame, B. M., C. Evans, and P. J. Roebber, 2017: The influence of PBL parameterization on the practical predictability of convection initiation during the Mesoscale Predictability Experiment (MPEX). *Wea. Forecasting*, **32**, 1161–1183.

Calhoun, R., R. Heap, M. Princevac, R. Newsom, H. Fernando, and D. Ligon, 2006: Virtual towers using coherent doppler lidar during the Joint Urban 2003 Dispersion Experiment. *J. Appl. Meteor. Climatol.*, **45**, 1161–1126.

Cambell, W. F., E. A. Satterfield, B. Ruston, and N. L. Baker, 2017: Accounting for correlated observation error in a dual-formulation 4D variational data assimilation system. *Mon. Wea. Rev.*, **145**, 1019–1032.

Carbone, R. E., J. D. Tuttle, D. A. Ahijevych, and S. B. Trier, 2002: Inferences of predictability associated with warm season precipitation episodes. *J. Atmos. Sci.*, **59**, 2033–2056.

Chen, I.-H., J.-S. Hong, Y.-T. Tsai, and C.-T. Fong, 2020: Improving afternoon thunderstorm prediction over taiwan through 3DVar-based radar and surface data assimilation. *Wea. Forecasting*, **35**, 2603–2620.

Chipilski, H. G., X. Wang, and D. B. Parsons, 2020: Impact of assimilating PECAN profilers on the prediction of bore-driven nocturnal convection: A multi-scale forecast evaluation for the 6 July 2015 case study. *Mon. Wea. Rev.*, **143**, 1147–1175.

Clark, R., 2016: FP3 Ellis, KS radiosonde data, version 2.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6GM85DZ.

Coniglio, M. C., S. F. Corfidi, and J. S. Kain, 2011: Environment and early evolution of the 8 May 2009 derecho-producing convective system. *Mon. Wea. Rev.*, **139**, 1083–1102.

Coniglio, M. C., S. J. Correia, 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.

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.

Corfidi, S. F., S. J. Corfidi, and D. M. Schultz, 2008: Elevated convection and castellanus: Ambiguities, significance, and questions. *Wea. Forecasting*, **23**, 1280–1303.

Crook, N. A., 1990: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785.

Crook, N. A., and J. B. Klemp, 2000: Lifting by convergence lines. *J. Atmos. Sci.*, **57**, 873–890.

Davis, C. A., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784.

Davis, C. A., K. W. Wanning, 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.

Degelia, S. K., X. Wang, and D. 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.

Degelia, S. K., X. Wang, D. Stensrud, and A. Johnson, 2018: Understanding the impact of radar and in situ observations on the prediction of a nocturnal convection initiation event on 25 June 2013 using an ensemble-based multiscale data assimilation system. *Mon. Wea. Rev.*, **143**, 3087–3108.

Degelia, S. K., X. Wang, D. Stensrud, 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.

Delgado, R., and K. Vermeesch, 2016: FP2 UMBC surface weather station data, version 1.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6SJ1HSG.

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.

Djalalova, I. V., and Coauthors, 2016: The POWER experiment: Impact of assimilation of a network of coastal wind profiling radars on simulating offshore winds in and above the wind turbine layer. *Wea. Forecasting*, **31**, 1071–1091.

Dowell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman Filter Experiments, journal=Mon. Wea. Rev., volume = 132, pages=1982-2005,.

Du, J., and Coauthors, 2014: NCEP regional ensemble update: Current systems and planned storm-scale ensembles. *26th Conf. on Weather Analysis and Forecasting/22nd Conf. on Numerical Weather Prediction*, Atlanta, GA, Amer. Meteor. Soc., J1.4, https://ams.confex.com/ams/94Annual/webprogram/Paper239 030.html.

Earth System Research Laboratory, 2016: The high-resolution rapid refresh (HRRR), rapidrefresh.noaa.gov/hrrr/.

Ecklund, W. L., D. A. Carter, and B. B. Balsley, 1988: A UHF wind profiler for the boundary layer: Brief description and initial results. *J. Atmos. Oceanic Technol.*, **5**, 432–441.

Ek, M., K. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Yano, and J. Tarpley, 2003: Implementation of Noah land surface model advances in the national centers for environmental prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851–8867.

Fast, J. D., and Coauthors, 2019: Overview of the HI-SCALE field campaign: A new perspective on shallow convective clouds. *Bull. Amer. Meteor. Soc.*, **100**, 821–840.

Feltz, W. F., W. L. Smoth, H. B. Howell, R. O. Knuteson, H. Woolf, and H. E. Revercomb, 2003: Near-continuous profiling of temperature, moisture, and atmospheric stability using the Atmospheric Emitted Radiance Interferometer (AERI). *J. Appl. Meteor.*, **42**, 584–597.

Fielding, M., and M. Janiskova, 2018: Assimilating profiles of cloud radar and lidar observations into the ECMWF 4D-Var system. *International Symposium on Data Assimilation 2018*, Munich, Germany, 3.6.

Fielding, M., and O. Stiller, 2019: Characterizing the representativity error of cloud profiling observations for data assimilation. *Atmospheres*, **124**, 4086–4103.

Fowler, A., and P. J. Van Leeuwen, 2013: Observation impact in data assimilation: the effect of non-gaussian observation error. *Tellus A: Dynamic Meteor. And Ocean.*, **65**, 20 035.

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

Geer, A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037.

Geer, A. J., and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1191–1217.

Geerts, B., and Coauthors, 2017: The 2015 plains elevated convection at night (PECAN) field project. *Bull. Amer. Meteor. Soc.*, **98**, 767–787.

Gero, J., H. Revercomb, D. D. Turner, J. Taylor, B. Ermold, K. Gaustad, R. Garcia, and D. Hackel, 2014: Atmospheric Emitted Radiance Interferometer (AERISUMMARY) from Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Atmospheric radiation measurement (ARM) climate research facility data archive, accessed 01 June 2019. doi:http://dx.doi.org/10.5439/10251465.

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.

Grell, G. A., and S. R. Freitas, 2013: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **13**, 23 845–23 893.

Gremillion, M. S., and R. E. Orville, 1999: Thunderstorm characteristics of cloud-to-ground lightning at the kennedy space center, florida: A study of lightning initiation signatures as indicated by the WSR-88D. *Wea. Forecasting*, **14**, 640–649.

Haghi, K. R., and Coauthors, 2018: Bore-ing into nocturnal convection. *Bull. Amer. Meteor. Soc.*, **100**, 1103–1121.

Hanesiak, J., and D. D. Turner, 2016a: FP3 University of Manitoba Doppler lidar wind profile data, version 1.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D60863P5.

Hanesiak, J., and D. D. Turner, 2016b: FP6 University of Manitoba Doppler lidar wind profile data, version 1.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D64F1NTN.

Hansell, R. A., and Coauthors, 2010: An assessment of the surface longwave direct radiative effect of airborne Saharan dust during the NAMMA field campaign. *J. Atmos. Sci.*, **67**, 1048–1065.

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.

Holdridge, D., and D. D. Turner, 2015: FP6 Hesston, KS radiosonde data, version 1.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6765CD0.

Hong, S., and J. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.

Horel, J., and Coauthors, 2002: Mesowest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211–225.

Horgan, K. L., D. M. Schultz, J. E. Hales, S. F. Corfidi, and R. H. Johns, 2007: A five-year climatology of elevated severe convective storms in the united states east of the rocky mountains. *Wea. Forecasting*, **22**, 1031–1044.

Houtekamer, P. L., and F. Zhang, 2016: Review of the ensemble kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532.

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.

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**, D13 103.

Janjić, T., and Coauthors, 2018: On the representation error in data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 1257–1278.

Johns, R. H., and C. A. Doswell, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612.

Johnson, A., and X. Wang, 2017: 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. *Wea. Forecasting*, **32**, 289–315.

Johnson, A., and X. Wang, 2019: Multicase assessment of the impacts of horizontal and vertical grid spacing, and turbulence closure model, on subkilometer-scale simulations of atmospheric bores during PECAN. *Mon. Wea. Rev.*, **147**, 1533–1555.

Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108.

Johnson, A., X. Wang, and S. K. Degelia, 2017a: Design and implementation of a GSI-based convection-allowing ensemble data assimilation and forecast system for the PECAN field experiment. Part II: Overview and evaluation of real-time system. *Wea. Forecasting*, **32**, 1227–1251.

Johnson, A., X. Wang, K. R. Haghi, and D. B. Parsons, 2017b: Evaluation of forecasts of a convectively generated bore using an intensively observed case study from PECAN. *Mon. Wea. Rev.*, **146**, 3097–3122.

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.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press.

Kawabata, T., H. Iwai, H. Seko, Y. Shoji, K. Saito, S. Ishii, and K. Mizutani, 2014: Cloud-resolving 4d-var assimilation of doppler wind lidar data on a meso-gamma-scale convective system. *Mon. Wea. Rev.*, **142**, 4484–4498.

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.

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.

Keene, R. M., and R. S. Schumacher, 2013: The bow and arrow mesoscale convective structure. *Mon. Wea. Rev.*, **141**, 1648–1672.

Klein, P., D. D. Turner, E. Smith, and J. Gebauer, 2016: Mobile PISA 1 OU/NSSL CLAMPS radiosonde data, version 1.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6416VDH.

Knupp, K., and R. Wade, 2016: MP2 UAH MIPS 915 MHz profiler NIMA-processed consensus wind and moments data, version 1.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6B27SJ2.

Lakshmanan, V., K. Hondl, and R. Rabin, 2009: An efficient, general-purpose technique for identifying storm cells in geospatial images. *J. Ocean. Atmos. Technol.*, **26**, 523–537.

Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The warning decision support system-integrated information. *Wea. Forecasting*, **22**, 596–612.

Lilly, D. K., 1990: Numerical prediction of thunderstorms—has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798.

Lin, Y., 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.

Loehrer, S. M., T. A. Edmands, and J. A. Moore, 1996: TOGA COARE upper-air sounding data archive: Development and quality control procedures. *Bull. Amer. Meteor. Soc.*, **77**, 2651–2672.

Martin, W. J., and M. Xue, 2006: Sensitivity analysis of convection of the 24 May 2002 IHOP case using very large ensembles. *Mon. Wea. Rev.*, **134**, 192–207.

Menzies, R. T., and R. M. Hardesty, 1989: Coherent doppler lidar for measurements of wind fields. *Proceedings of the IEEE*, **77**, 449–462.

Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky satellite radiance. *Mon. Wea. Rev.*, **145**, 1063–1081.

Morse, C. S., R. K. Goodrich, and L. B. Cornman, 2002: The NIMA method for improved moment estimation from doppler spectra. *J. Atmos. Oceanic Technol.*, **19**, 274–295.

Nakanishi, M., and H. Niino, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407.

National Research Council, 2009: Observing weather and climate from the ground up. National Academies Press, Washington, D.C., http://www.nap.edu/catalog/12540.

Newsom, R., and R. Krishnamurthy, 2014: Doppler Lidar (DLAUX) from Southern Great Plains (SGP) Central Facility, Lamont, OK (C1). Atmospheric radiation measurement (ARM) climate research facility data archive, accessed 01 June 2019. doi:http://dx.doi.org/10.5439/1374838.

NOAA, 2015: Warning Decision Training Division (WDTD) Virtual Lab: MRMS products guide (version 10). Accessed 1 July 2020, https://vlab.ncep.noaa.gov/web/wdtd/-/reflectivity-at- x-1?selectedFolder=562123.

O'Connor, E. J., A. J. Illingworth, I. M. Brooks, C. D. Westbrook, R. J. Hogan, F. Davis, and B. J. Brooks, 2010: A method for estimating the turbulent kinetic energy dissipation rate from a vertically pointing doppler lidar, and independent evaluation from balloon-borne in situ measurements. *J. Atmos. Oceanic Technol.*, **27**, 1652–1664.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, **56**, 527–530.

Peters, J. M., E. R. Nielsen, M. D. Parker, S. M. Hitchcock, and R. S. Schumacher, 2017: The impact of low-level moisture errors on model forecasts of an MCS observed during PECAN. *Mon. Wea. Rev.*, **145**, 3599–3624.

Privé, N. C., R. M. Errico, and K.-S. Tai, 2017: The impacts of increased frequency of rawinsonde observations on forecast skill investigated with an observing system simulation experiment. *Mon. Wea. Rev.*, **142**, 1823–1834.

Reif, D. W., and H. B. Bluestein, 2017: A 20-year climatology of nocturnal convection initiation over the central and southern Great Plains during the warm season. *Mon. Wea. Rev.*, **145**, 1615–1639.

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.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608.

Rogers, E., and Coauthors, 2009: NCEP north american mesoscale modeling system: Recent changes and future plans. *19th Conf. on Numerical Weather Prediction and 23rd Conf. on Weather Analysis and Forecasting*, Omaha, NE, Amer. Meteor. Soc., 2A.4.

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

Schwartz, C. S., and R. A. Sobash, 2017: Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations. *Mon. Wea. Rev.*, **145**, 3397–3418.

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.

Shao, H., and Coauthors, 2016: Bridging research to operations transitions: Status and plans of community GSI. *Bull. Amer. Meteor. Soc.*, **97**, 1427–1440.

Shin, H., and S.-Y. Hong, 2011: Intercomparison of planetary boundary-layer parameterizations in the WRF model for a single day from CASES-99. *Bound.-Layer. Meteor.*, **139**, 261–281.

Sisterson, D. L., R. A. Peppler, T. S. Cress, P. J. Lamb, and D. D. Turner, 2016: The arm southern great plains (sgp) site. *Meteorological Monographs*, **57**, 6.1–6.14.

Sivaraman, C., L. Ma, L. Riihimaki, P. Muradyan, R. Coulter, S. Collis, and S. Xie, 1990: Radar wind profiler (915RWPPRECIPCON; updated hourly) from Southern Great Plains (SGP) central facility (C1), NW radar wind profiler site (I10), NE radar wind profiler site (8). Atmospheric radiation measurement (ARM) climate research facility data archive, accessed 01 June 2018. doi:https://dx.doi.org/10.5439/1025127.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR, 125 pp.

Skinner, P. S., L. J. Wicker, D. M. Wheatley, and K. H. Knopfmeier, 2016: Application of two spatial verification methods to ensemble forecasts of low-level rotation. *Wea. Forecasting*, **31**, 713–735.

Smalikho, I., F. Köpp, and S. Rahm, 2005: Measurement of atmospheric turbulence by 2-μm Doppler Lidar. *J. Atmos. Oceanic Technol*, **22**, 1733–1747.

Smith, E., J. A. Gibbs, E. Fedorovich, and T. Bonin, 2015: WRF model study of the great plains low-level jet: Effects of grid spacing and boundary layer parameterizations. *22nd Symposium on Boundary Layers and Turbulence, Amer. Meteor. Soc.*, June, 2016.

Smith, E. N., G. J. Gebauer, P. M. Klein, E. Fedorovich, and J. A. Gibbs, 2019: The Great Plains low-level jet during PECAN: Observed and simulated characteristics. *Mon. Wea. Rev.*, **147**, 1845–1869.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630.

Smith, W. L., W. F. Feltz, R. O. Knuteson, H. E. Revercomb, H. M. Woolf, and H. B. Howell, 1999: The retrieval of planetary boundary layer structure using ground-based infrared spectral radiance measurements. *J. Atmos. Oceanic Technol*, **16**, 323–333.

Sobash, R. A., and D. J. Stensrud, 2015: Assimilating surface mesonet observations with the EnKF to improve ensemble forecasts of convection initiation on 29 May 2012. *Mon. Wea. Rev.*, **143**, 3700–3725.

Stalker, J., J. Lasley, G. Frederick, R. McPherson, P. Campbell, B. Philips, and B. Pasken, 2013: A nationwide network of networks. *Bull. Amer. Meteor. Soc.*, **94**, 1602–1606.

Stelten, S., and W. A. Gallus, 2017: Pristine nocturnal convective initiation: A climatology and preliminary examination of predictability. *Wea. Forecasting*, **32**, 1613–1635.

Stewart, L. M., S. L. Dance, N. K. Nichols, J. R. Eyre, and J. Cameron, 2013: Estimating interchannel observation-error correlations for IASI radiance data in the Met Office system. *Quart. J. Roy. Meteor. Soc.*, **140**, 1236–1244.

Storm, B., J. Dudhia, S. Basu, A. Swift, and I. Giammanco, 2009: Evaluation of the weather research and forecasting model on forecasting low-level jets: Implications for wind energy. *Wind Energy*, **12**, 81–90.

Sun, J., and S. B. Trier, 2018: Physical processes leading to elevated convection initiation during 25-26 June PECAN: Convective-scale reanalysis based on a radar data assimilation system. *A Special Symposium on Plains Elevated Convection At Night (PECAN)*, Amer. Meteor. Soc., 1.6, https://ams.confex.com/ams/98Annual/webprogram/Paper336 167.html.

Sun, J., S. B. Trier, Q. Xiao, M. L. Weisman, H. Wang, Z. Ying, M. Xu, and Y. Zhang, 2012: Sensitivity of 0–12-h warm-season precipitation forecasts over the central United States to model initialization. *Wea. Forecasting*, **27**, 832–855.

Surcel, M., M. Berenguer, and I. Zawadzki, 2010: The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part I: Methodology and seasonal comparison. *Mon. Wea. Rev.*, **138**, 3084–3106.

Tandeo, P., P. Ailliot, A. Carrassi, T. Miyoshi, M. Pulido, and Y. Zhen, 2018: A review of innovation-based methods to jointly estimate model and observation error covariance matrices in ensemble data assimilation. *Mon. Wea. Rev.*, **148**, 3973–3994.

Tao, W.-K., and Coauthors, 2003: Microphysics, radiation and surface processes in the Goddard Cumulus Ensemble (GCE) model. *Meteor. Atmos. Phys.*, **82**, 97–137.

Toms, B. A., J. M. Tomaszewski, D. D. Turner, and S. E. Koch, 2017: Analysis of a lower-tropospheric gravity wave train using direct and remote sensing measurement systems. *Mon. Wea. Rev.*, **145**, 2791–2812.

Trier, S. B., S. D. Kehler, and J. Hanesiak, 2020: Observations and simulation of elevated nocturnal convection initiation on 24 June 2015 during PECAN. *Mon. Wea. Rev.*, **148**, 613–635.

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.

Trier, S. B., R. D. Roberts, J. Sun, T. M. Weckwerth, and J. W. Wilson, 2018: Physical processes influencing elevated convection initiation during 25-26 June PECAN: observations and numerical simulations. *A Special Symposium on Plains Elevated Convection At Night (PECAN)*, Amer. Meteor. Soc., 1.5, https://ams.confex.com/ams/98Annual/webprogram/Paper335 614.html.

Trier, S. B., J. W. Wilson, D. A. Ahijevych, and R. A. Sobash, 2017: Mesoscale vertical motions near nocturnal convection initiation in PECAN. *Mon. Wea. Rev.*, **145**, 2919–2941.

Turner, D. D., 2016a: FP2 AERIoe thermodynamic profile retrieval data, version 2.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/d6x63k9k.

Turner, D. D., 2016b: FP3 AERIoe thermodynamic profile retrieval data, version 2.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6Z31WV0.

Turner, D. D., 2016c: FP5 AERIoe thermodynamic profile retrieval data, version 2.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D61V5C5J.

Turner, D. D., 2016d: FP6 ARM surface meteorology data, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6RR1WN0.

Turner, D. D., 2016e: MP1 OU/NSSL CLAMPS Doppler lidar VAD wind data, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6BR8QJH0.

Turner, D. D., 2016f: MP1 OU/NSSL CLAMPS Doppler lidar VAD wind data, version 1.2. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2019. doi:https://doi.org/10.5065/D6VQ312C.

Turner, D. D., and W. G. Blumberg, 2019: Improvements to the AERIoe thermodynamic profile retrieval algorithm. *IEEE Select Topics Appl. Earth Obs. Remote Sens.*, **12**, 1339–1354.

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.

UCAR/NCAR, 2015a: FP1 ARM central facility radiosonde data, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://data.eol.ucar.edu/dataset/485.021.

UCAR/NCAR, 2015b: FP3 FP4 FP5 QC 5 min surface data, tilt corrected, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6BZ645V.

UCAR/NCAR, 2015c: FP4 NCAR/EOL 915 MHz profiler NIMA consensus winds and moments, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6RV0KXH.

UCAR/NCAR, 2015d: FP5 NCAR/EOL 915 MHz profiler 30 minute consensus winds and moments data, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6H993DQ.

UCAR/NCAR, 2016a: FP4 NCAR/EOL QC soundings, version 2.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D63776XH.

UCAR/NCAR, 2016b: FP5 NCAR/EOL QC soundings, version 2.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6ZG6QF7.

UCAR/NCAR, 2016c: MP4 NCAR/EOL MISS 915 MHz profiler 30 minute consensus winds and moments and surface meteorology data, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6RJ4GPJ.

UCAR/NCAR, 2016d: MP4 NCAR/EOL QC soundings, version 2.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6707ZNV.

Vaisala, 2017: Vaisala Radiosonde RS41 Measurement Performance. Ref. B211356EN-B, 28 pp, https://www.vaisala.com/sites/default/files/documents/WEA-MET-RS41-Performance- White-paper-B211356EN-B-LOW-v3.pdf.

Vermeesch, K., 2015: FP2 Greensburg, KS radiosonde data, version 1.0. UCAR/NCAR – Earth Observation Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/d6fq9tph.

Wagner, T., D. D. Turner, and R. Newsom, 2016a: MP3 University of Wisconsin SPARC AERIoe thermodynamic profile data, version 1.0. UCAR/NCAR – Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/d60z71hc.

Wagner, T., D. D. Turner, and R. Newsom, 2016b: MP3 University of Wisconsin SPARC Doppler lidar VAD wind data, version 2.0. UCAR/NCAR - Earth Observing Laboratory., accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6V9869Bh.

Wagner, T., D. D. Turner, and R. Newsom, 2016c: MP3 University of Wisconsin SPARC radiosonde data, version 2.0. UCAR/NCAR - Earth Observing Laboratory., accessed 01 June 2018. doi:https://dx.doi.org/10.5065/d6vh5m7b.

Wagner, T., D. D. Turner, and R. Newsom, 2016d: MP3 University of Wisconsin SPARC surface meteorological data, version 1.0. UCAR/NCAR - Earth Observing Laboratory., accessed 01 June 2018. doi:https://dx.doi.org/10.5065/d6n014xz.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008b: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments. *Mon. Wea. Rev.*, **136**, 5132–5147.

Wang, X., T. M. Hamill, J. S. Whitaker, and C. H. Bishop, 2009: A comparison of the Hybrid and EnSRF analysis schemes in the presence of model errors due to unresolved scales. *Mon. Wea. Rev.*, **137**, 3219–3232.

Wang, X., and T. Lei, 2014: GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. *Mon. Wea. Rev.*, **142**, 3303–3325.

Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble-variational hybrid data assimilation for NCEP Global Forecast System: single resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117.

Wang, X., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222–227.

Wang, Y., and X. Wang, 2017: Direct assimilation of radar reflectivity without tangent linear and adjoint of the nonlinear observation operator in the GSI-based EnVar system: Methodology and experiment with the 8 May 2003 Oklahoma City tornadic supercell. *Mon. Wea. Rev.*, **145**, 1447–1471.

Waugh, S., and C. Ziegler, 2017: NSSL mobile mesonet data, version 1.1. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D64M92RG.

Weckwerth, T. M., J. Hanesiak, J. W. Wilson, S. B. Trier, S. K. Degelia, W. A. Gallus, R. D. Roberts, and X. Wang, 2019: Nocturnal convection initiation during PECAN 2015. *Bull. Amer. Meteor. Soc.*, **100**, 2223–2239.

Weckwerth, T. M., and D. B. Parsons, 2006: A review of convection initiation and motivation for IHOP$_2$002.*Mon. Wea. Rev.*,**134**,5 − −22.

Weckwerth, T. M., and U. Romatschke, 2019: Where, when, and why did it rain during PECAN? *Mon. Wea. Rev.*, **147**, 3557–3573.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60A**, 62–79.

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.

Weston, P., W. Bell, and J. Eyre, 2014: Accounting for correlated error in the assimilation of high-resolution sounder data. *Tellus*, **60A**, 62–79.

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.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

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.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. Academic Press.

Wilson, J. W., and R. D. Roberts, 2006: Summary of convective storm initiation and evolution during IHOP: Observational and modeling perspective. *Mon. Wea. Rev.*, **134**, 23–47.

Wilson, J. W., S. B. Trier, D. W. Reif, R. D. Roberts, and T. M. Weckwerth, 2018: Nocturnal elevated convection initiation of the PECAN 4 July hailstorm. *Mon. Wea. Rev.*, **146**, 243–262.

Wulfmeyer, V., H. 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$_2$002.*Mon. Wea. Rev.*,**134**,209 − −230.

Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638.

Zhu, K., Y. Pan, M. Xue, X. Wang, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A regional GSI-based ensemble Kalman filter data assimilation system for the rapid refresh configuration: Testing at reduced resolution. *Mon. Wea. Rev.*, **141**, 4118–4139.

Ziegler, C., M. Coniglio, M. Parker, and R. Schumacher, 2016: CSU/NCSU/NSSL MGAUS radiosonde data, version 3.0. UCAR/NCAR - Earth Observing Laboratory, accessed 01 June 2018. doi:https://dx.doi.org/10.5065/D6W66HXN.