THE INFLUENCE OF VERTICAL ADVECTION DISCRETIZATION IN WRF-ARW MODEL ON CAPPING INVERSION REPRESENTATION IN WARM-SEASON, THUNDERSTORM SUPPORTING ENVIRONMENTS

by

David Nevius

A Thesis Submitted in

Partial Requirement of the

Requirements for the Degree of

Master of Science

in Atmospheric Science

at

The University of Wisconsin – Milwaukee

May 2018

ABSTRACT

THE INFLUENCE OF VERTICAL ADVECTION DISCRETIZATION IN WRF-ARW MODEL ON CAPPING INVERSION REPRESENTATION IN WARM-SEASON, THUNDERSTORM SUPPORTING ENVIRONMENTS
by

David Nevius

The University of Wisconsin-Milwaukee, 2018
Under the Supervision of Professor Clark Evans

This study evaluates forecasts of capping inversions and thermodynamic variables for believed areas of possible deep, moist convection initiation during the warm-season using the Weather Research and Forecasting Model (WRF) with the Advanced Research core (WRF-ARW). WRF-ARW was configured nearly identical to the National Severe Storms Laboratory (NSSL) version of WRF (NSSL-WRF). WRF-ARW's default third-order-accurate vertical advection scheme, which is an odd-order-accurate scheme, is known to introduce implicit damping which acts to dampen short wavelength features (Skamarock et al. 2008), such as capping inversions. It is hypothesized that by increasing WRF-ARW's vertical advection to the next higher, even-order-accurate vertical advection scheme, this would remove the associated implicit dampening, thus improving WRF-ARW's handling of capping inversion representation. After computing Student's *t* tests on the bias of each weather and thermodynamic variable, it was deemed that the fourth-order-accurate vertical advection scheme did not improve WRF-ARW's representation of capping inversion or other thermodynamic variables. Despite the rejected hypothesis, this study does confirm that the Mellor-Yamada-Janjić (MYJ) planetary boundary layer (PBL) parametrization has a cool and moist bias near the surface, as also found by Coniglio et al. (2013), Burlingame et al. (2017), Cohen et al. (2015) , Clark et al. (2012), among others. It is

likely that the poor representation of capping inversions in WRF-ARW is from other numerics in the model, which is beyond the scope of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I first want to thank my advisor, Dr. Clark Evans, for his mentorship, knowledge, and assistance during my time at UWM. My experience at UWM has been a positive one, in large part because of Dr. Evans. I would also like to thank Dr. Jon Kahl and Dr. Paul Roebber for their input and willingness to be on my committee. A big thanks to the other Atmospheric Science graduate students for the great laughs, trips to Chick-fil-A, and help with numerous programing issues. I greatly appreciate UWM's excellent HPC facilities for much of the computing power for this research, along with the University of Wyoming's Department of Atmospheric Science's upper air data. Last but not least, I would like to thank my family and friends for much support.

1. **Introduction**

Thermodynamic parameters such as convective inhibition (CIN) and convective available potential energy (CAPE) are important predictors for thunderstorm development and help forecasters gauge atmospheric stability. Previous studies have shown that models using the Advanced Research dynamical core of the Weather Research and Forecasting (WRF-ARW; Skamarock et al. 2008, Powers et al. 2017) model have limitations that hinder their ability to accurately predict these parameters in thunderstorm-supporting environments (e.g., Burlingame et al. 2017, Coniglio et al. 2013, Jirak et al. 2015, Cohen et al. 2015, 2017, Kain et al. 2016), thus limiting numerical forecast skill. Since WRF-ARW is used by NOAA as the basis for its current-generation convection-allowing forecast system (e.g., Benjamin et al. 2016), these limitations can impact both deterministic and probabilistic forecasts of convective phenomena. The governing motivation behind this research is to determine if altering the vertical advection formulation in WRF-ARW improves thermodynamic profiles, with a specific focus on capping and subsidence inversion representation in warm-season, thunderstorm-supporting environments.

Each spring as part of the Hazardous Weather Testbed (HWT), NOAA conducts the Spring Forecasting Experiment (SFE; e.g., Kain et al. 2003, Clark et al. 2012), working with the National Weather Service (NWS) and National Severe Storms Laboratory (NSSL) to promote collaboration between research and operations while evaluating new technologies and science for NWS operations. During the 2015 HWT SFE (Gallo et al. 2017), participants were asked the following question: "*Compare forecast soundings in regions with elevated mixed layers (EMLs) from the NSSL-WRF [e.g., Coffer et al. 2013] and 2.2 km Operational UM (Unified Model [Walters et al. 2014, Wood et al. 2014]) at sites where observed raob [rawinsonde observation] data is available. With a focus on sounding structure in the PBL and depiction of any capping*

inversions, which model has the best forecast sounding?" From the 89 responses, 67% answered the UM was better than the NSSL-WRF, 10% answered that the UM was worse than the NSSL-WRF, and 21% stated they performed about the same (Jirak et al. 2015). Therefore, one of the preliminary findings from the SFE 2015 was that strong vertical gradients in temperature and moisture associated with capping inversions were better resolved in the UK Met Office convection-allowing Unified Model compared to the NSSL-WRF (Jirak et al. 2015, Gallo et al. 2017). Qualitatively, similar findings were obtained from an analogous evaluation conducted during the 2014 SFE: the NSSL-WRF tended to have a smoothed representation of capping inversions compared to the UM (Fig. 1; Kain et al. 2016). This significantly impacts the vertical distribution of negative buoyancy for ascending near-surface-based parcels, which may play a crucial role in forecasting the timing, location, and incidence of convection initiation (e.g., Trier et al. 2014). One hypothesis tested to potentially better represent thermodynamic profiles in thunderstorm-supporting environments was to increase the model's vertical resolution, but early testing showed that this was not sufficient to resolve this issue (Kain et al. 2016, Burlingame et al. 2017).

Most capping or subsidence inversions (e.g., Lanicci and Warner 1991, Farrell and Carlson 1989) occur immediately atop the PBL. Because the temporal and spatial scales of turbulent vertical mixing in the PBL are finer than those resolved by a numerical model, a parameterization scheme is needed to simulate the micrometeorology within the PBL. PBL schemes aid in simulating subgrid-scale physical processes; however, these schemes can produce errors, thus impacting atmospheric variables within and near the PBL (e.g., Coniglio et al. 2013 and references therein). To that end, Coniglio et al. (2013) quantified sensitivity in model-derived vertical profiles within the PBL to the choice of PBL parameterization within short-

range, warm-season, convection-allowing WRF-ARW model forecasts. Whenever a capping inversion was observed, model-derived mixed-layer convective inhibition (MLCIN) was typically under-forecast and the sharp vertical temperature and moisture gradients associated with the inversion were often damped considerably relative to observations regardless of the PBL scheme used (e.g., Fig. 2). Similar findings for capping inversion representation were obtained by Burlingame et al. (2017) in ensemble forecasts of convection initiation events during the 2013 Mesoscale Predictability Experiment (Weisman et al. 2015, Trapp et al. 2016). Therefore, it can be surmised that the PBL parameterization alone in WRF-ARW is not the sole or primary contributor to poorly modeled capping inversion representations.

In WRF-ARW, finite difference approximations are used to calculate partial derivatives within the governing equations. The default formulations use odd-order-accurate formulations for horizontal (fifth-order) and vertical (third-order) advection (Skamarock et al. 2008), which introduce implicit numerical damping of short wavelength features (e.g., Fig. 1 of Wicker and Skamarock 2002) such as capping inversions. These odd-order formulations are comprised of the next higher (even) ordered scheme plus a residual term that acts as an implicit diffusion operator of the same order as the next-highest-order-accurate differencing scheme (Skamarock et al. 2008). In contrast, the even-order-accurate formulations do not include implicit damping and have been shown to be slightly more accurate than their next-lower-order-accurate formulation for simple one-dimensional advection examples (Wicker and Skamarock 2002). However, the even-order-accurate formulations have stricter numerical stability criteria and are numerically dispersive.

In contrast to WRF-ARW, the UM uses a semi-Lagrangian formulation to solve the non-hydrostatic, compressible deep-atmosphere equations of motion, which is not associated with

implicit numerical damping for advection terms (Walters et al. 2014, Wood et al. 2014). It is hypothesized that the improved capping inversion representation in the UM relative to WRF-ARW results from this absence of implicit damping. Additionally, a primary contributor to the development and, in particular, maintenance of capping inversions is from large-scale subsidence. Based on the aforementioned minimal sensitivity in WRF-ARW-modeled capping inversion representation to vertical grid spacing and PBL parameterization, it is hypothesized that the implicit damping of short wavelengths associated with the odd-order-accurate vertical advection formulation in WRF-ARW is the primary cause of degraded capping inversion representation.

One potential way to improve capping inversion representation in WRF-ARW would be to solve the equations of motion using a semi-Lagrangian formulation, as the UM does. However, this would be quite an undertaking, and would birth an entirely new model. Another approach would be to adjust the vertical advection finite differencing formulation used by WRF-ARW from an odd-order-accurate formulation to an even-order-accurate formulation, which would remove the implicit damping (Wicker and Skamarock 2002). However, the loss of implicit damping comes at a cost, as using an even-order-accurate spatial finite differencing formulation with the third-order Runge-Kutta time integration scheme used by WRF-ARW results in stricter numerical stability criteria (Table 1) and numerical dispersion, particularly of shorter wavelength features (Wicker and Skamarock 2002). This study uses a fourth-order-accurate vertical advection formulation in WRF-ARW and evaluates whether the fourth-order-accurate formulation better represents vertical thermodynamic profiles, including capping inversions, than the default third-order-accurate formulation. We hypothesize the even-order accurate formulation to better represent capping inversions because of its reduced damping of short wavelength

features as tied to vertical advection and subsidence, each of which may be found in environments where surface-based and mixed-layer parcels contain negative buoyancy between their lifting condensation level and level of free convection. Despite the stricter stability criteria, we hypothesize that it will not be necessary to alter the time step to maintain numerical stability. The WRF-ARW-recommended time step is $6\Delta x$ (for Δx in km; Skamarock et al. 2008), and with this time step and 35 vertical levels, the maximum-allowed values of vertical velocity to ensure numerical stability remain reasonable (e.g., $w_{max} \leq$ 25-30 m s^{-1}) for both even- and odd-ordered-accurate schemes.

The structure of this paper is as follows. Section 2 describes the model configuration, sounding identification and evaluation used for this work. Section 3 discusses the key results of the temperature and moisture profiles, in addition to the derived thermodynamic variables. A summary and discussion of the results from and key implications of the research are provided in section 4.

2. **Methods**

a. Model description

This study uses WRF-ARW version 3.8.1 to conduct model simulations for an approximate one-month period concurrent with the 2017 SFE (3 – 31 May), with the exception of 15, 17, 19, and 21 May due to simulation failures, using third- and fourth-order-accurate finite difference formulations for vertical advection. The NSSL-WRF configuration (Coffer et al. 2013) of WRF-ARW is used in this study as it has been in place and largely unchanged for nearly a decade and provides a benchmark for convection-allowing forecasts. This configuration includes using the Mellor-Yamada-Janjić (MYJ) PBL parameterization (Janjić 1994), WRF single-moment 6-class microphysics scheme (Hong et al. 2006), Rapid Radiative Transfer Model (RRTM) for longwave

radiation (Mlawer et al. 1997), Dudhia shortwave radiation (Dudhia et al. 1989), Noah land-surface model (Tewari et al. 2004), positive-definite advection of moisture (Skamarock and Weisman 2009), a 4-km horizontal grid length over a 1200 x 800 conterminous United States domain (Fig. 3), 35 vertical levels (including 10 in the lowest 1-2 km), a time step of 24 s, and a forecast length of 36 h. The diffusion operator term was set to coordinate surfaces, and the eddy coefficient operator was set just for horizontal deformation, thus any vertical diffusion is done by the PBL scheme. Additionally, the sixth-order numerical diffusion operator was turned off. Initial and lateral boundary conditions are obtained from 40-km North American Mesoscale model data obtained from EMC/NCEP. Model data are post-processed using wrf-python (Ladwig 2017). The control simulation which uses the NSSL-WRF's default third-order accurate vertical advection scheme is hereafter called the control simulation. The second simulation using the fourth-order accurate vertical advection scheme is hereafter called the fourth-order simulation.

b. Sounding Identification

The primary goal of this study is to evaluate model performance in forecasting vertical thermodynamic profiles, particularly when capping inversions are present, in areas where deep, moist convection is believed to be possible. For verification, routine NWS radiosonde observations are considered to be the best-available "truth." Only radiosonde locations that are within SPC's 0600 UTC Day 1 Convective Outlook (valid for the subsequent 1200 UTC to 1200 UTC period) for a given event are considered at forecast lead times of 11 h (1100 UTC) and 23 h (2300 UTC), which are the hours closest to actual radiosonde release times.

For example, verifying 3 May forecasts only include radiosonde locations within SPC's 3 May 0600 UTC Day 1 Convective Outlook for the period 1200 UTC 3 May – 1200 UTC 4 May

2017 (Fig. 4). For the 11-h forecast time, nocturnal processes such as the low-level jet, a residual layer, and radiational cooling at the surface are often present. Additionally, for nocturnal PBLs, the primary contributor to turbulent vertical mixing is vertical wind shear. For the 23-h forecast, both buoyancy and mechanically driven turbulent vertical mixing and its effects on PBL thermodynamic properties are found. All observed and model vertical profiles are linearly interpolated to a common grid with an interval of 10 hPa. It would be desireable to interpolate to a finer resolution to better resolve potential caps, however, due to the relatively coarse vertical resolution of the model simulations, 10 hPa was chosen as the desired interval to balance both the model resolution and the finer observed sounding resolution.

c. Sounding Evaluation

As discussed above, 1665 soundings were identified during our period from within SPC's Day 1 Convective Outlook. When verifying model-derived thermodynamic profiles, it is important to exclude any soundings that were contaminated by precipitation, in order to gain an accurate representation of the convection supporting environment. To filter out contaminated soundings, a similar method to, but a subset of, that of Coniglio et al. (2013) is followed. For observed rainfall, precipitation data are gathered from NCEP's Stage IV (Lin and Mitchell 2005) hourly multi-sensor precipitation analysis for the continental United States. A sounding location is not used in the evaluation process if there is observed rainfall or simulated rainfall, from either simulation, of greater than or equal to 0.5 mm hourly accumulation at any grid point within 40 km in any one hour during the 3 h prior to the sounding.

The procedure of removing contaminated soundings produces a total of 809 soundings from the original set, which are then subset based on the presence of a capping inversion and, if one is present, cap strength. To distinguish between soundings with and without capping inversions, a

subset of the criteria first outlined by Farrell and Carlson (1989) is used. Within a subsidence layer or EML, temperature and dew point profiles typically follow a dry adiabat and mixing ratio isopleth, respectively. At the base of this layer, there is typically notable warming and drying. Our approach to identifying the presence of a cap is to identify this small layer of warming and drying at the base of the subsidence or EML. To do so, it is first determined if temperature increases by any amount within any 20 hPa layer (e.g., over three interpolated vertical grid points) between the surface to 600 hPa. Limiting the upper bound to 600 hPa ensures that the tropopause is not captured by this methodology over the range of events considered. Next, it is determined if dew point temperature decreases by at least 2°C over that same 20 hPa layer. Including this criterion eliminates surface-based inversions caused by nocturnal radiational cooling.

For soundings that have been identified to have a capping inversion, cap strength is determined by calculating the parcel buoyancy minimum (B_{min}), or the minimum value of the surface-based lifted parcel's temperature minus the environment's temperature (Trier et al. 2014). A weak cap is defined to have B_{min} between 0°C and -2°C and a strong cap is defined for $B_{min} \leq$ -2°C. These thresholds are chosen following Granziano and Carlson (1987), who found that values of what they dubbed lid strength (defined nearly equivalently to B_{min}) of magnitude 2°C and greater effectively identify whether or not surface-based convection initiation will occur.

To illustrate the effectiveness of the inversion identification method, observed soundings from 18 May 2017 are depicted in Fig. 5. The sounding from Dallas-Fort Worth, TX (KFWD) depicts a strong capping inversion (B_{min} = -6°C; Fig. 5b). Large negative buoyancy is observed between ~825 hPa and ~690 hPa, with a rapid increase in temperature and rapid decrease in dew

point temperature with height noted at approximately 850 hPa. The sounding from Slidell, LA (KLIX) depicts a weak capping inversion (B_{min} = -1.5°C; Fig. 5c). At approximately 750 hPa, a substantial decrease in dew point temperature and a small increase in temperature are observed. Finally, the sounding from Davenport, IA (KDVN) contains no capping inversion (Fig. 5d). Although there is a layer around 810 hPa where temperature increases with height, there is an insufficient decrease in dew point temperature with height within this layer to qualify as a capping inversion. Because of variability between soundings, it is difficult to objectively and accurately identify every sounding with (and exclude those without) a capping inversion using this or any methodology. However, subjective analysis of inversion classifications over the samples considered herein suggest that the method utilized in this study effectively and accurately identifies most soundings with a capping inversion of any strength and nearly all soundings with well-defined capping inversions. After the applying the capping inversion identification algorithm to the set of soundings not contaminated by precipitation, there are 383 soundings with capping inversions, of which 43 are classified as weak and 340 are classified as strong. Of the uncontaminated soundings with a cap present, 275 occurred at the 11-h lead time and the remaining 108 occurred at the 23-h lead time. Several additional soundings were removed due to missing data caused by malfunctions during the radiosonde launch.

The *MetPy* (May et al. 2017) package is used to calculate all thermodynamic variables from both the observed and model sounding profiles, including lifted condensation level (LCL), level of free convection (LFC), equilibrium level (EL), B_{min}, mixed-layer CAPE and CIN (MLCAPE, MLCIN; here, a 100-hPa-deep mixed layer is assumed), most-unstable CAPE and CIN (MUCAPE, MUCIN), and surface-based CAPE and CIN (SBCAPE, SBCIN). It is worth noting that *MetPy* does not use the virtual temperature correction when evaluating thermodynamic

parameters; however, the internally consistent method used for both modeled and observed soundings suggests that the quantative analysis is likely to be robust. The example soundings presented in Trier et al. (2014) generally include an LCL, single LFC, and single EL at or near the tropopause. However, soundings may have multiple LFCs and thus multiple ELs, and B_{min} might be more negative for the second rather than first negative buoyancy layer. As formulated, *MetPy* does not look beyond the first LFC, which is not a comprehensive calculation, and only computes CIN for any lifted parcel through the first EL. In light of this, B_{min} is defined as the buoyancy minimum between the surface-based parcel and the environment over any layer of negative buoyancy below a sounding's final EL.

For all uncontaminated soundings within areas of possible deep-moist convection, including the sample subsets of soundings with and without capping inversions present, aggregate statistics are computed using *SciPy's* statistics package for the 11-h and 23-h forecast lead times. For vertical profiles of temperature, potential temperature, dew point temperature, and mixing ratio, mean bias and mean absolute error are computed for the lowest four km above ground level (AGL) of each sounding. Because the samples of each weather variable and thermodynamic variable are approximately Gaussian at all altitudes (not shown; computed using normaltest from *SciPy's* statistics package), a standard Student's *t* test for means of two independent samples, with an appropriate Fligner-Killeen test for equal variance between the two data sets, is used to determine whether or not the null hypothesis that the mean error between the control simulation and the fourth-order simulation for a particular variable is equal to zero can be rejected. The full sample size for computing the Student's *t* test is used as opposed to a smaller effective sample size, similar to Coniglio et al. (2013), which determined that there is an autocorrelation between individual soundings in the sample. The effect of using a smaller sample size would be to

decrease the likelihood that differences are statistically significant, which as we discuss in the results, is not a concern due to the lack of statistical significance between the two samples.

In this study, we present confidence (%), which is calculated as $100 \times (1 - p)$, where p is the two-tailed decimal probability of the hypothesis being false. In other words, the lower the p value, the higher the confidence in rejecting the null hypothesis of there being statistically significant difference between the two model simulations. The null hypothesis is rejected, and the differences between the control and fourth-order simulation are said to be statistically significant, if and only if the confidence exceeds 95%.

3. Results

a. Temperature profiles

Mean bias and mean absolute error profiles for temperature and potential temperature forecasts between 0-4 km AGL are shown in Fig. 6a,b and Fig. 7a,b, respectively. For the 11-h (morning) forecast, the mean bias of temperature in the lowest 4 km is fairly close to zero for both samples. For potential temperature during the morning forecast, there is a slight warm bias for the entire displayed profile. Coniglio et al. (2013) also identifies a similar warm bias for the 11-h forecast, which they hypothesize is introduced from the initial conditions. For our case, the NAM has undergone multiple updates over the past several years since Coniglio et al. (2013); thus, the same hypothesis cannot be made without further investigation.

For the 23-h (late afternoon) forecast, there is a noticeable cold bias in the lowest ~0.75 km in both temperature and potential temperature for both samples (Figs. 6-7b). Just above this level, there is a warm bias, particularly noticeable in the potential temperature profile from about 0.75 to 2.0 km (Fig. 7b). It has been shown in previous studies (e.g., Coniglio et al. 2013, Burlingame et al. 2017, Cohen et al. 2015, Clark et al. 2015) that the MYJ boundary layer

scheme has a tendency to undermix, especially in daytime convective boundary layers. This change in sign of the bias across the boundary is likely representing too little mixing, resulting in a cool bias near the ground and a warm bias in the upper boundary layer.

The statistical confidence in the differences in bias for both morning and late afternoon temperature and potential temperature forecasts is too low at all altitudes for the null hypothesis of non-zero differences to be rejected.

b. Moisture profiles

Dew point temperature (Fig. 6c) has a mostly warm bias during the morning forecast, with mean absolute error increasing with height, for both the control and fourth-order samples. It is worth mentioning the spread of the distribution for dew point temperature is much larger compared to distribution of temperature for both forecast hours. Even though the mean bias is near zero for most of the morning and late afternoon profiles within each sample, the large spread indicates that errors in dew point temperature are highly variable between individual soundings.

In conjunction with the temperature profiles showing that the MYJ scheme often produces too little mixing, the morning and especially late afternoon mixing ratio profiles provide further supporting evidence for this assertion. For the lowest ~0.75 km in both the 11-h and 23-h forecasts for water vapor mixing ratio (Figs. 6-7d), there is a moist bias with the exception of a small layer near the surface. Just above this layer, a dry bias exists in the morning samples from ~0.75 km to 1.50 km and from ~0.75 km to 1.75 km in the late afternoon samples. The late afternoon samples of dew point temperature also follow the same trend (Fig 7c).

As with the temperature and potential temperature samples at both forecast lead times, the confidence in the differences in mean absolute error being non-zero between the control and

fourth-order samples for both dew point temperature and water vapor mixing ratio is not statistically significant.

c. Derived sounding variables

Similar to the temperature and moisture profile results, the difference in thermodynamic variables listed in Section 2 also do not support our original hypothesis. Mixed-layer parcels in the morning for both samples tend to be more stable than observed, with MLCAPE mean errors of ~ -500 J kg^{-1} and MLCIN errors ~ 80 J kg^{-1}, which we hypothesize is likely from errors from within the PBL and radiation parameterizations. Note that CAPE and CIN are only computed if both the models and the observed soundings have positive buoyancy. Coniglio et al. (2013) also identify a low bias in model-forecast MLCAPE and MLCIN for the 11-h forecast that is insensitive to the PBL parameterization used. For the 23-h forecast, the mean errors of MLCAPE and MLCIN for both samples are centered around 0 J kg^{-1} .

As with the vertical temperature and moisture profiles, there is no statistically significant improvement in MLCAPE or MLCIN forecasts for the fourth-order simulations relative to the control, whether for all cases or only cases with a cap present, at both forecast lead times (not shown).

For most-unstable parcels (Figs. 10-11), the morning and late afternoon forecasts tend to be slightly less stable than observations for both the control and fourth-order samples, with MUCIN mean errors of ~ -25 J kg^{-1} and MUCAPE mean errors near 0 J kg^{-1} for both the morning and late afternoon forecasts. For surface-based parcels (Figs. 12-13), mean error for all samples during the morning and late afternoon is near 0 J kg^{-1}.

As with MLCAPE and MLCIN, as well as vertical temperature and moisture profiles, there is no statically significant improvement in the fourth-order simulation compared to the control for both MUCAPE and MUCIN as well as SBCAPE and SBCIN (not shown).

Finally, B_{min} magnitude and height errors are shown in Figs. 14-15. Recall that B_{min} is defined as the minimum of the surface-based parcel minus environment temperature, thus B_{min} is a negative value in °C when CIN is non-zero. With this is mind, positive errors in B_{min} indicate where a given model sample underpredicts, in its sample mean, negative buoyancy. In contrast, negative errors indicate where the model sample overpredicts, in its sample mean, negative buoyancy. Both the control and fourth-order samples underpredict negative buoyancy in the morning, with mean errors of ~0.75 to 1°C (Fig. 14). The mean error in height of B_{min} is centered at 0 hPa for both simulations. For the late afternoon forecast (Fig. 15), both the control and fourth-order samples again underpredict negative buoyancy, with only a marginal reduction in bias with the fourth-order simulation compared to the control. However, this reduction is not statistically significant for either forecast hour (not shown). Note also that there is a smaller mean error for the samples that include all cases versus the samples that only include capping inversions (Fig. 15), implying that both the control and fourth-order simulations underpredict cap strength, with no improvement from the fourth-order sample for cases with capping inversions.

Despite the underprediction of cap strength, both the control and fourth-order samples are associated with a deeper negative buoyancy layer. Specifically, the mean errors of SBLCL are ~15 to 25 hPa in the afternoon forecast for both samples (Fig. 16), or lower to the ground, whereas mean surface-based LFC (SBLFC) errors are centered around 0 hPa for both samples (Fig. 17). In an idealized sounding, the CIN layer is found between the LCL and LFC. With a lower SBLCL and unchanged SBLFC, on average, this indicates a deeper layer of CIN in the

simulations; this, in turn, results in near-zero bias, on average, for predictions of SBCIN from both the control and fourth-order simulations.

4. Conclusions and Discussion

This study implements a fourth-order-accurate vertical advection scheme in WRF-ARW and compares thermodynamic profiles and variables in regions where deep, moist convection initiation is believed to be possible to WRF-ARW's default third-order-accurate vertical advection scheme. Odd-order-accurate schemes introduce implicit damping which are known to dampen short wavelength features (Skamarock et al. 2008), such as capping inversions. It is hypothesized that by increasing WRF-ARW's vertical advection to the next higher, even-order-accurate vertical advection scheme, this would remove the associated implicit dampening, thus improving WRF-ARW's handling of capping inversion representation.

During the one month of forecasts in May 2017, a total of 1665 sounding locations are identified for model verification at 11-h and 23-h lead times. Precipitation contamination filtering results in a reduction to 493 soundings in the 11-h forecast and 316 soundings in 23-h forecast. Soundings are further subset by identifying if there a capping inversion exists and, if so, capping inversion strength, with 383 soundings with a capping inversion, including 275 at 11-h and 108 at 23-h. Derived thermodynamic variables are computed for each sounding, and mean bias and mean absolute error are computed for both the control and fourth-order simulation samples relative to observations at each forecast lead time. A Student's *t* test is performed to assess whether differences in bias between the control and fourth-order samples for each weather and thermodynamic variable are statistically significant.

There is no statistical significance in any differences between the control and fourth-order simulation samples for any variable and level considered, and thus no improvement to WRF-

ARW's representation of capping inversions through use of the fourth-order-accurate vertical advection formulation. Although the hypothesis motivating this research is rejected, this study does corroborate previous studies in showing that the MYJ PBL scheme has a tendency to mix too little in daytime, convective boundary layers, resulting in conditions that are too cool and moist near the surface. Additionally, WRF-ARW has a tendency to underpredict the cap strength, but in many cases maintains a similar amount of CIN by having a deeper negative buoyancy layer in the vertical. A representative sounding profile for observations and both control and fourth-order simulations from Dallas-Fort Worth, Texas at 27 May 2017 1100 UTC (Fig. 18) helps visualize this argument. At ~860 hPa, a well-defined subsidence inversion is present in observations (observed B_{min} = -7.36°C). Table 2 provides the thermodynamic variables for each profile, and both model simulations underpredict cap strength by about 1.5°C. Both simulations, including the fourth-order simulation, smooth the sharp increase in temperature right near ~860 hPa. Despite the simulations underpredicting the strength of the cap and smoothing the sharp gradient of temperature, the amount of CIN from both simulations is fairly close to the observed. This is in part due to the increased depth in the negative buoyancy layer. Additionally, near the surface, the models tend to be too moist and cool compared to the observed.

In light of the previous work of increasing vertical resolution, using various PBL parameterization schemes, and other model numerics, as described in the introduction, in addition to the results from this study, it is likely that the issue of accurately resolving capping inversion in WRF-ARW is a combination of these topics, and the specific details will likely vary between models. For instance, the Global Forecast System Model (GFS) uses semi-Lagrangian numerics, similar to the UM, but is perhaps worse than WRF-ARW in representing capping inversions. In the case of the GFS, it uses somewhat coarse vertical grid spacing, relative to

WRF-ARW and the UM in the boundary layer, and the GFS also uses a PBL parameterization that is one of the strongest vertical mixers under convectively unstable environments.

5. Tables

3rd Order	4th Order	5th Order	6th Order
1.61	1.26	1.42	1.08

Table 1. Maximum stable Courant number for one-dimensional linear advection when using the Runge-Kutta 3rd-order-accurate temporal differencing scheme (from Wicker and Skamarock 2002).

	Observed	Control	Fourth-Order
MLCAPE [J kg^{-1}]	3330.60	4373.79	4051.20
MLCIN [J kg^{-1}]	335.12	269.32	299.75
MUCAPE [J kg^{-1}]	3062.44	4373.79	4111.16
MUCIN [J kg^{-1}]	385.86	269.32	293.90
SBCAPE [J kg^{-1}]	3062.44	4373.79	4051.20
SBCIN [J kg^{-1}]	385.86	269.34	299.75
B_{min} [°C]	-7.36	-5.57	-5.79
B_{min} Height [hPa]	860	850	820
SBLCL [hPa]	921.99	962.30	955.44
SBLFC [hPa]	673.44	955.69	951.68
SBEL [hPa]	164.13	153.73	154.39

Table 2. *Thermodynamic variables for each profile found in Fig. 18.*

6. Figures

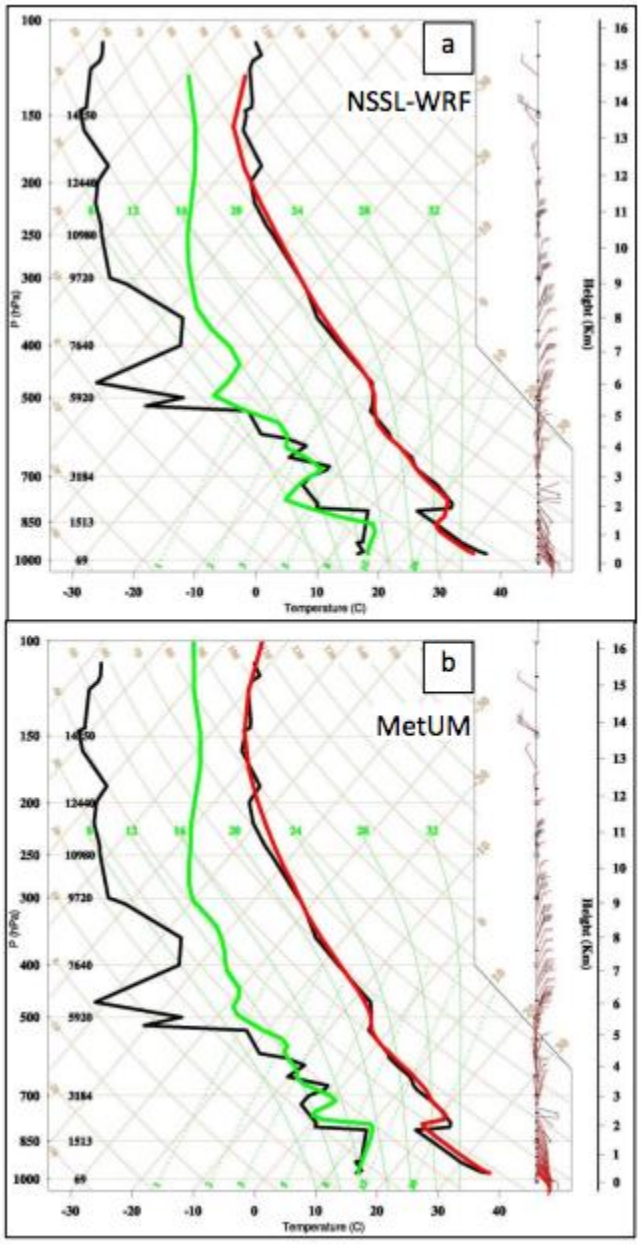

Figure 1. *An example showing corresponding point forecast soundings (24-h forecasts; temperature in °C in red, dew point temperature in °C in green) for the a) NSSL-WRF and b) MetUM, overlaid on observations (thick black profiles, in °C). Figure reproduced from Kain et al. (2016), their Fig. 9.*

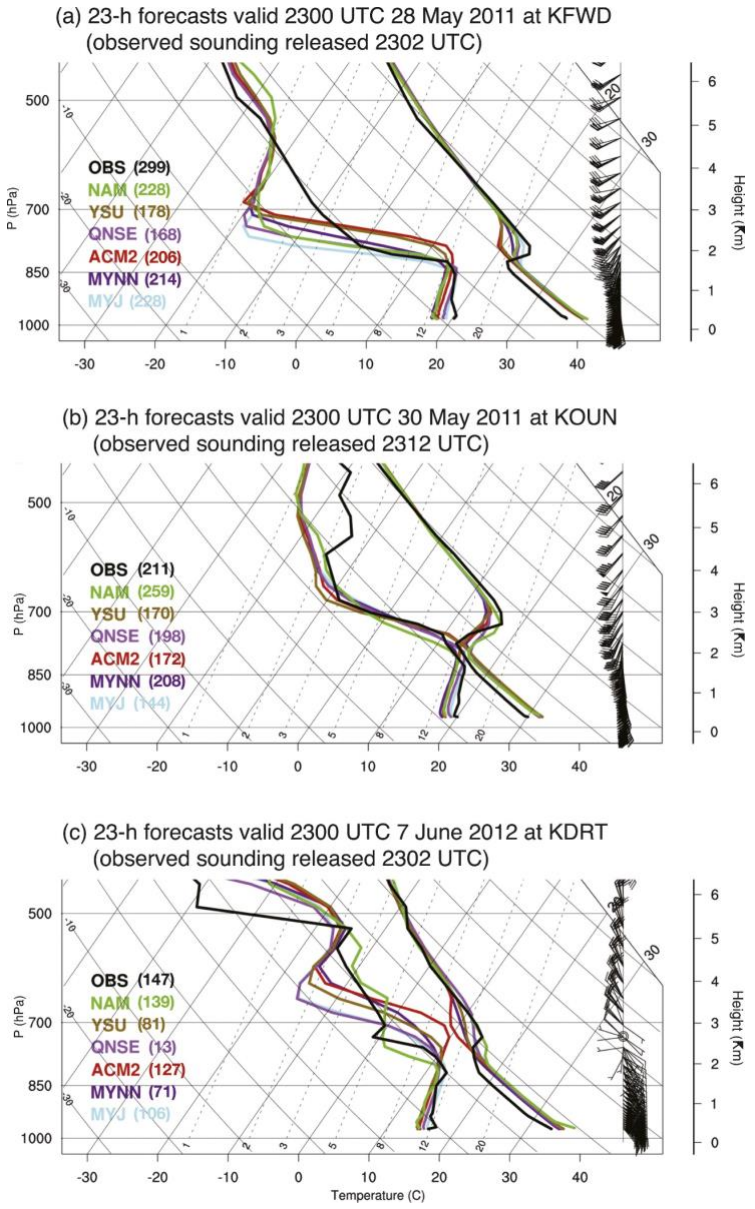

Figure 2. *Skew T–ln p diagrams showing examples of the models' smoothed representation of a capping inversion. Each colored line represents a model run with the respective PBL scheme used, and the black line is the observed sounding. Absolute magnitudes of MLCIN (J kg^{-1}) for each sounding are given in parentheses: 23-h forecasts valid at 2300 UTC (a) 28 May at Fort Worth, TX (KFWD), (b) 30 May at Norman, OK (KOUN), and 7 June at Del Rio, TX (KDRT), all in the year 2011. Reproduced from Coniglio et al. (2013), their Fig. 15.*

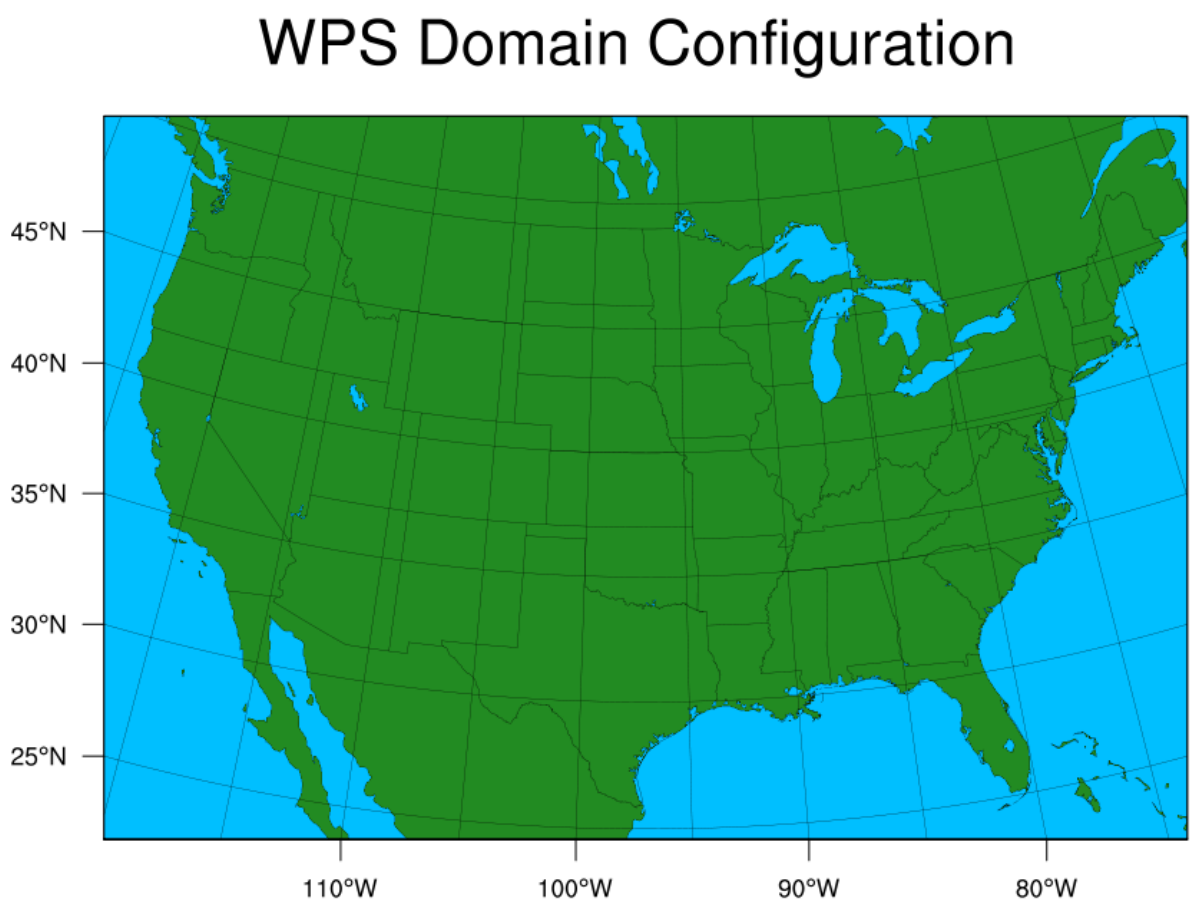

Figure 3. *The domain used for all numerical model simulations, encompassing the conterminous United States.*

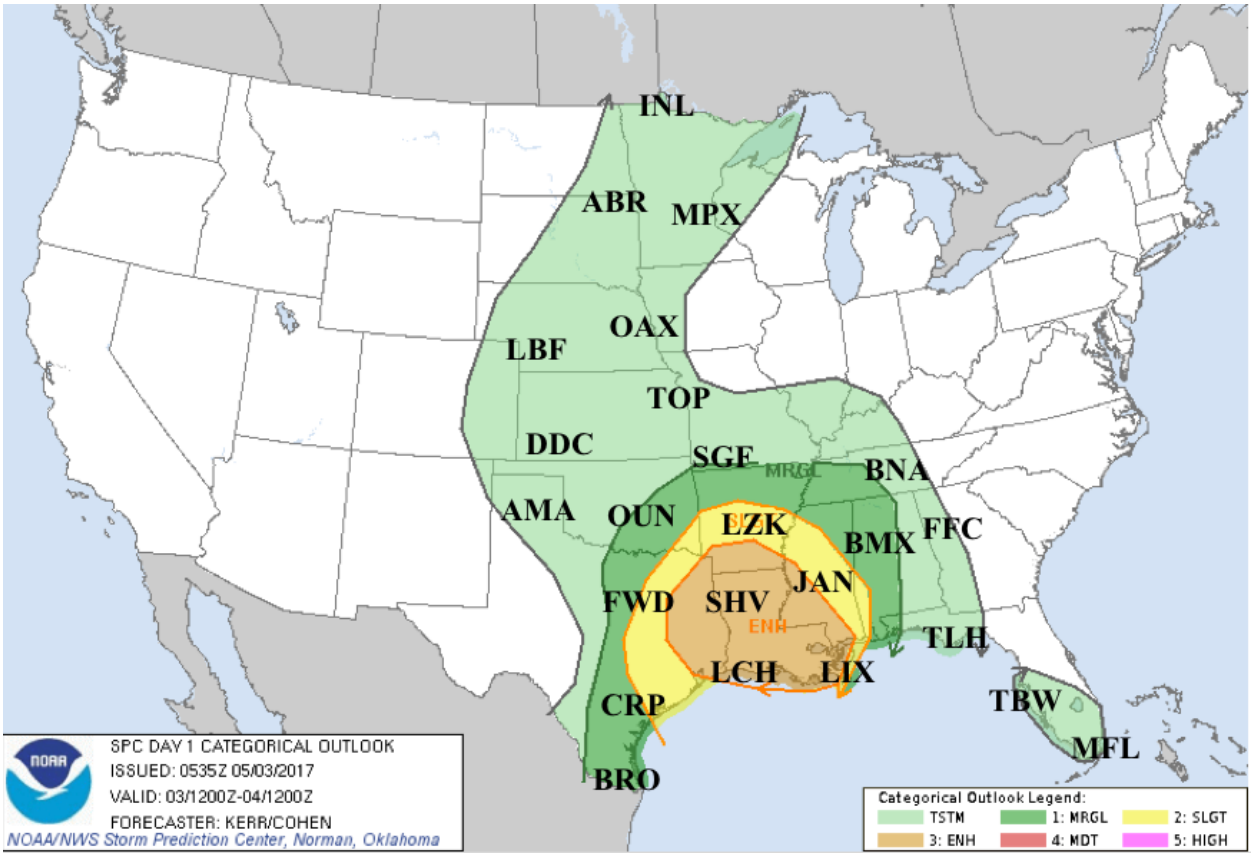

Figure 4. *SPC Day 1 Convective Outlook issued at 0535 UTC 3 May 2017 for the period 1200 UTC 3 May to 1200 UTC 4 May 2017 (shaded per the legend at lower right; where TSTM = Thunderstorm, MRGL = Marginal, SLGT = Slight, ENH = Enhanced, MDT = Moderate, and HIGH = High risk areas), and the radiosonde locations that are verified for the 11 and 23 h lead l-times for this case.*

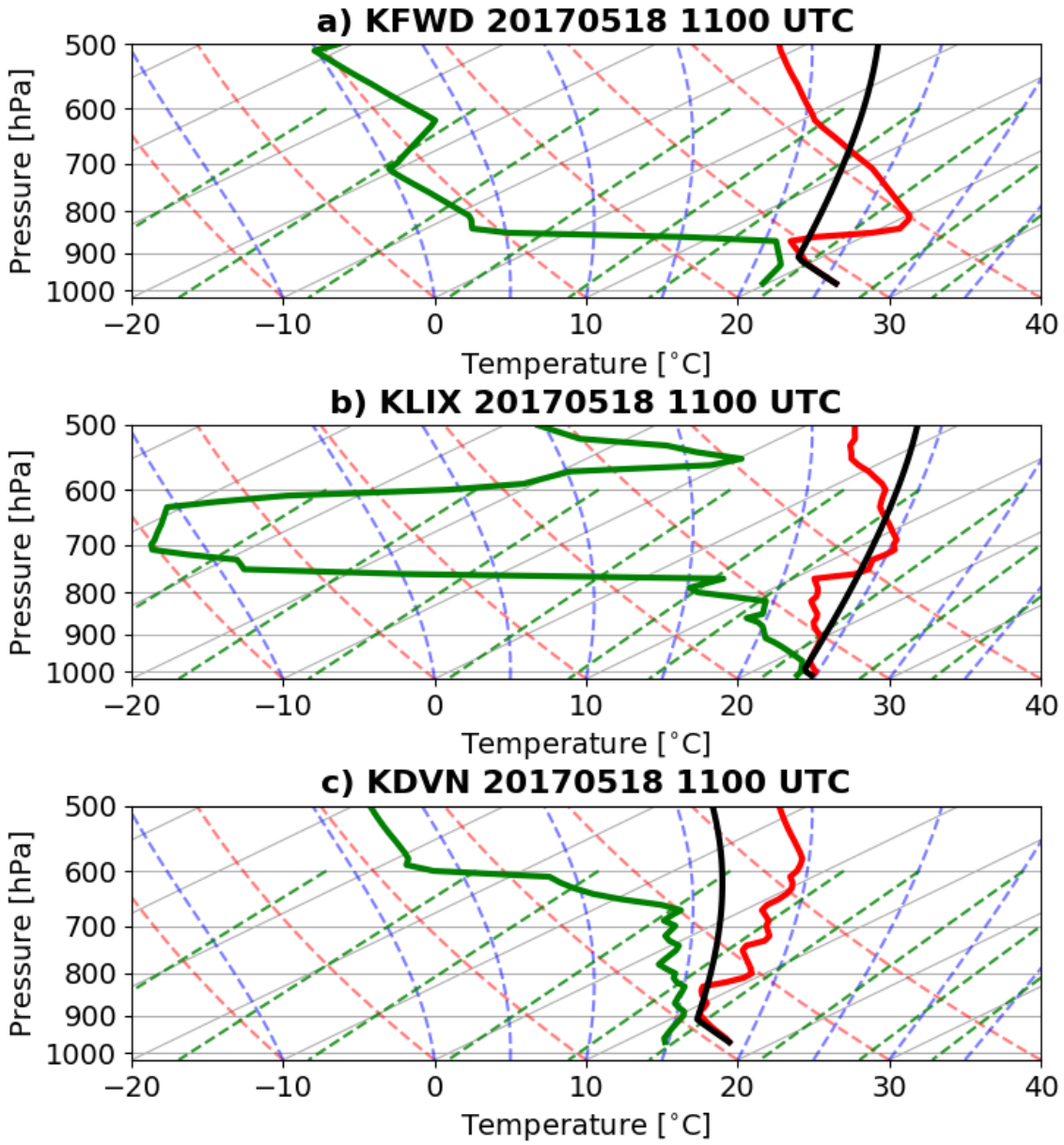

Figure 5. *Observed Dallas-Fort Worth, TX (KFWD; a), Slidell, LA (KLIX; b), and Davenport, IA (KDVN; c) skew T-ln p diagrams (red line: temperature in °C; green line: dew point temperature in °C; black line: parcel ascent curve for a surface-based parcel) valid at 1100 UTC 18 May 2017.*

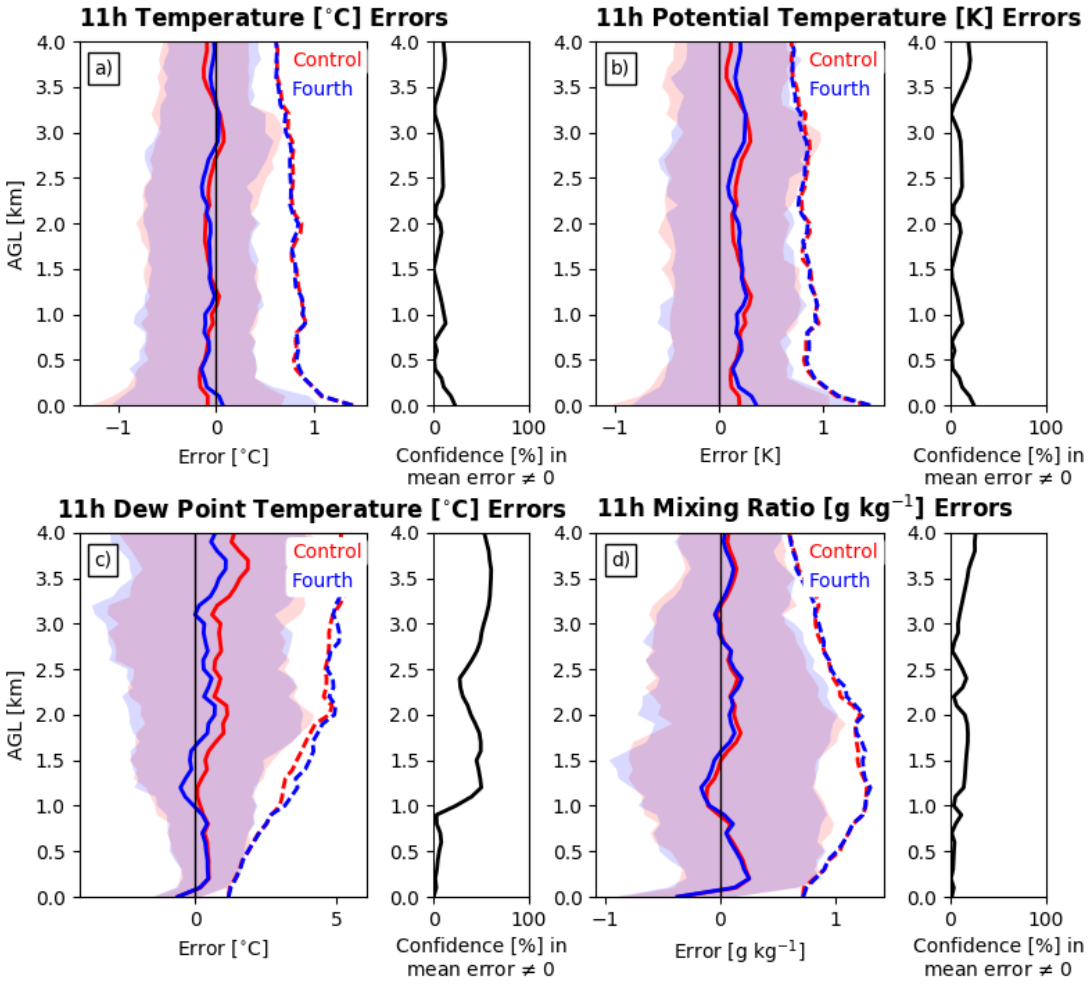

Figure 6. *In (a), (left) bias (solid) and mean absolute error (dashed) for the control (red) and fourth-order (blue) samples for the 11-h forecast temperature (°C), from the surface to 4 km AGL. Shading represents the distribution between the 25th and 75th percentiles. (right) Vertical profile of confidence in the bias difference being non-zero between the control and fourth-order samples. Panels (b), (c), and (d) are analogous to (a) but for potential temperature (K), dew point temperature (°C), and mixing ratio (g kg^{-1}), respectively.*

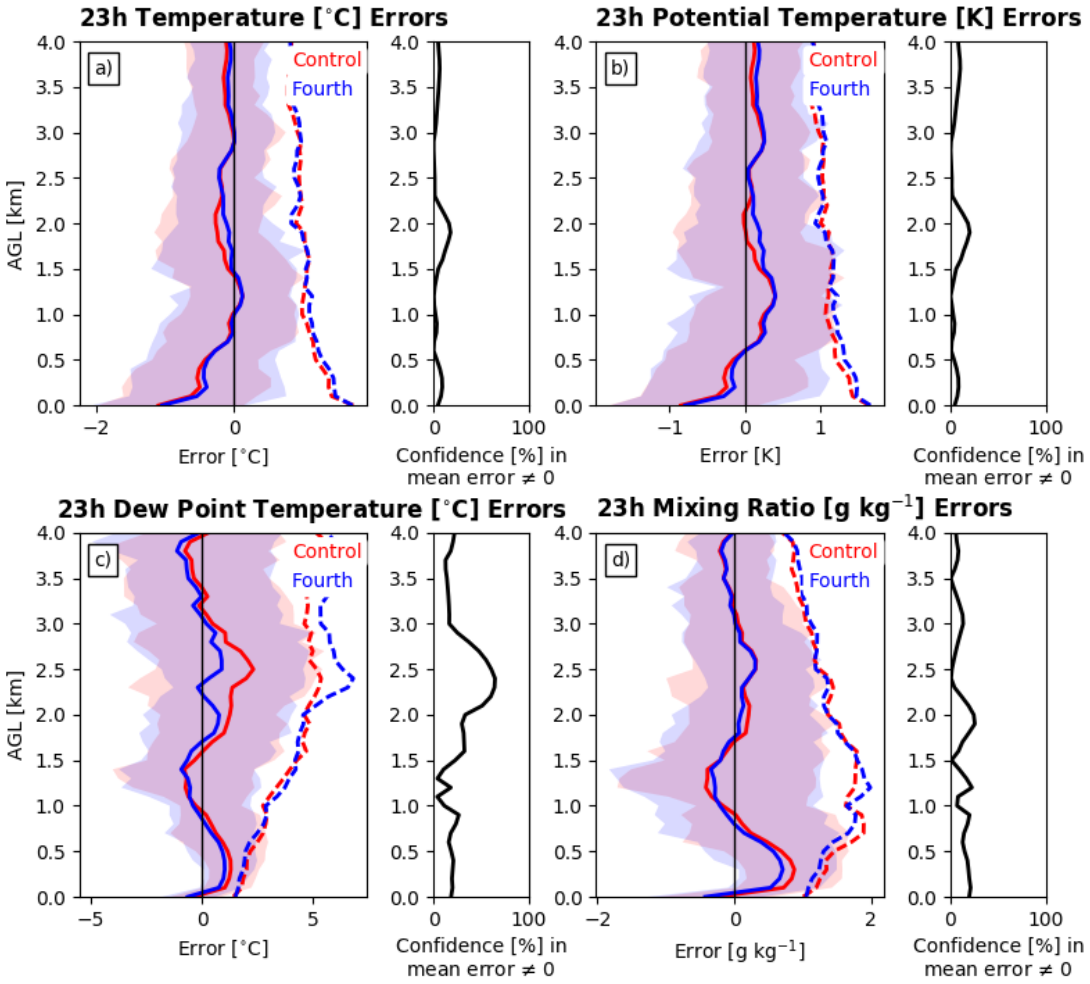

Figure 7. *As in Fig. 6, except for the 23-h forecast.*

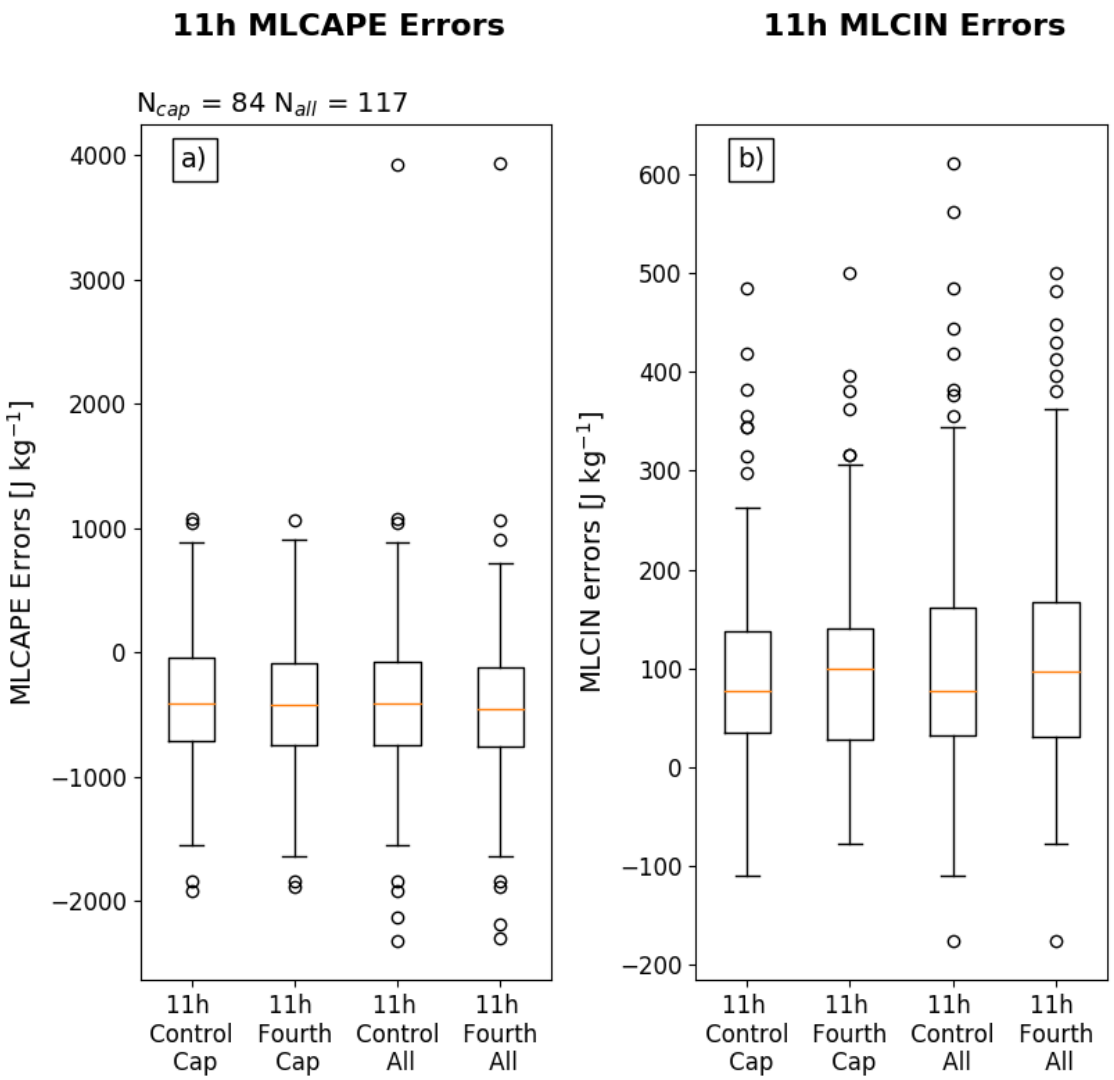

Figure 8. *Box-and-whisker diagrams of forecast MLCAPE (left) and MLCIN (right) for the 11-h forecast lead time for the control (Control) and fourth-order (Fourth) samples. The center orange line represents the mean of the sample, with the box enclosing the distribution between the 25th and 75th percentiles. The whiskers represent the maximum and minimum values, excluding outliers denoted in circles. The cases which denote 'cap' represent the sample which includes capping inversions (weak and strong) and cases that denote 'all' represent the sample of all soundings (i.e., soundings with a capping inversion as well as soundings with no identified capping inversion). In this study, CIN is assigned as a positive number. Cases with zero positive buoyancy are removed from all samples.*

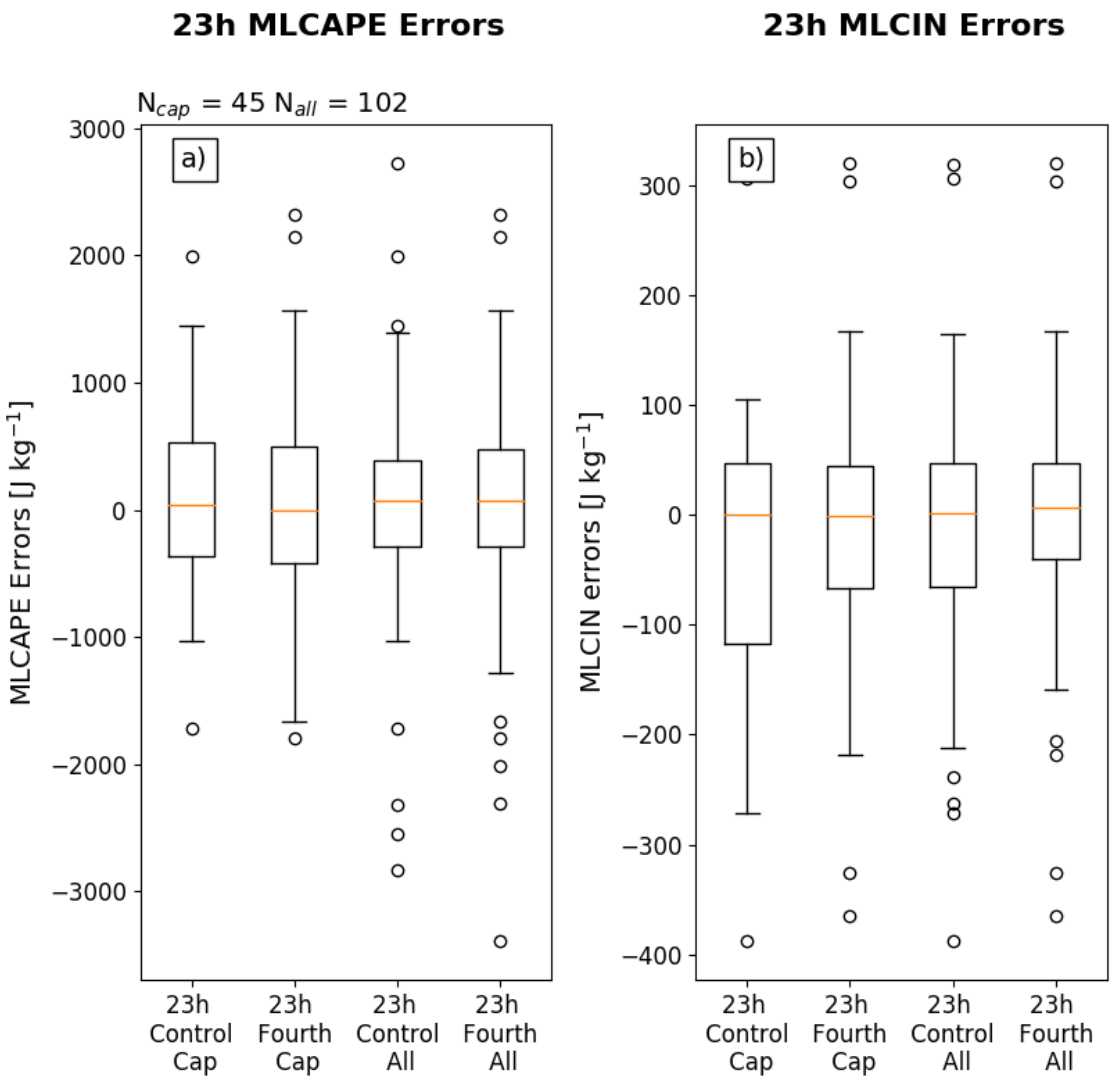

Figure 9. *As in Fig. 8, but for the 23-h forecast.*

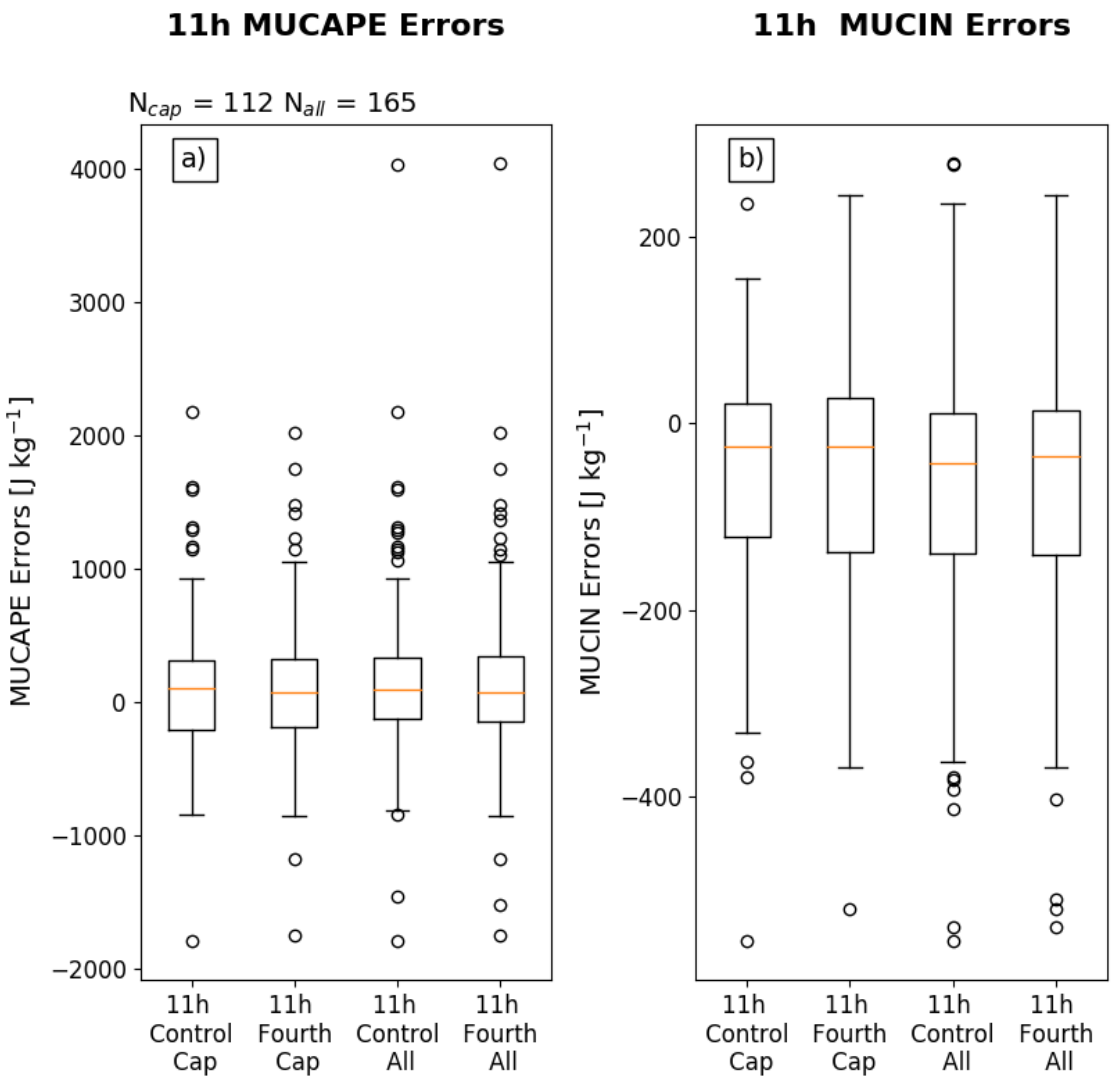

Figure 10. *As in Fig. 8, but for MUCAPE and MUCIN for the 11-h forecast.*

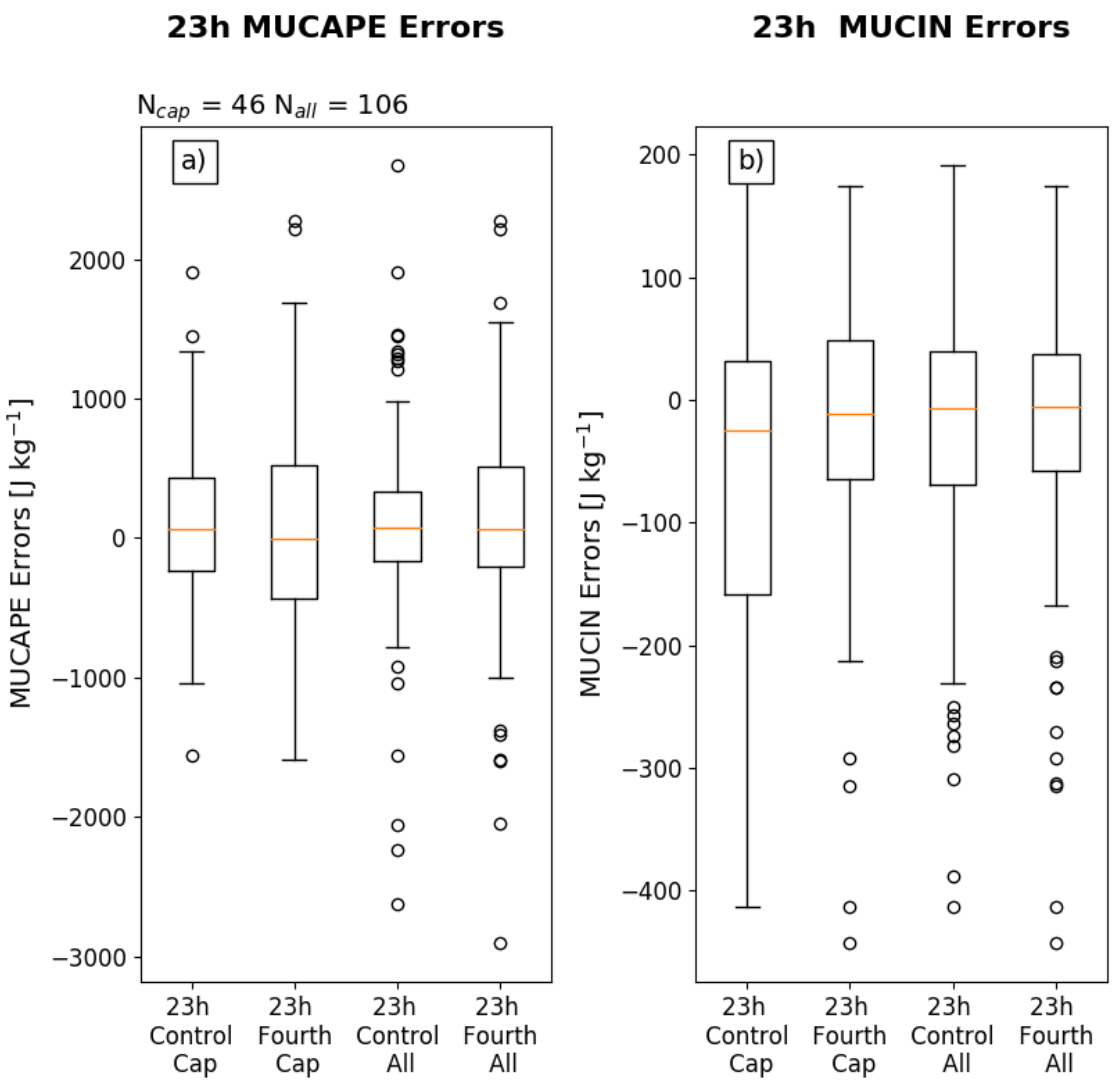

Figure 11. *As in Fig. 8, but for MUCAPE and MUCIN for the 23-h forecast.*

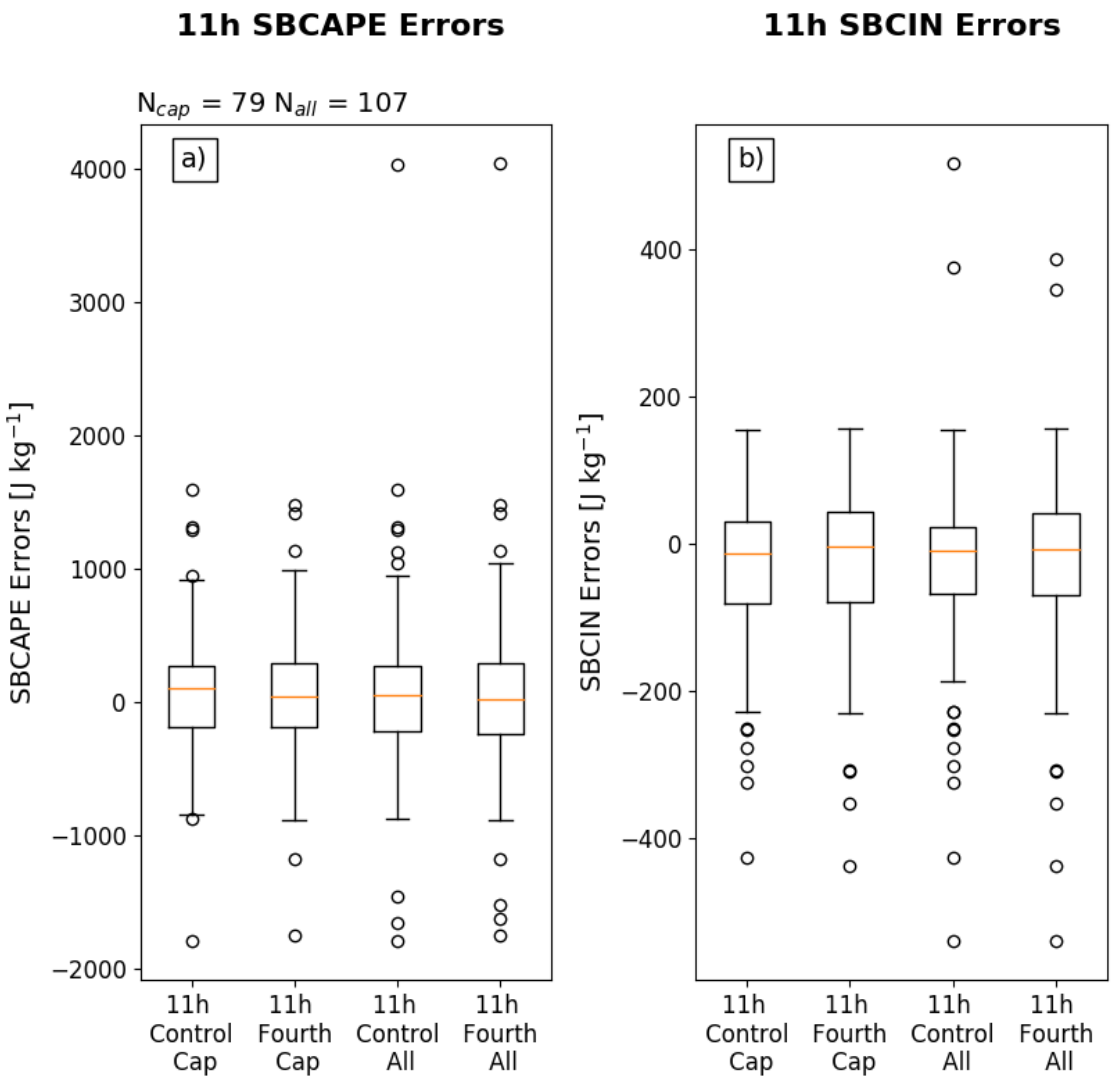

Figure 12. *As in Fig. 8, but for SBCAPE and SBCIN for the 11-h forecast.*

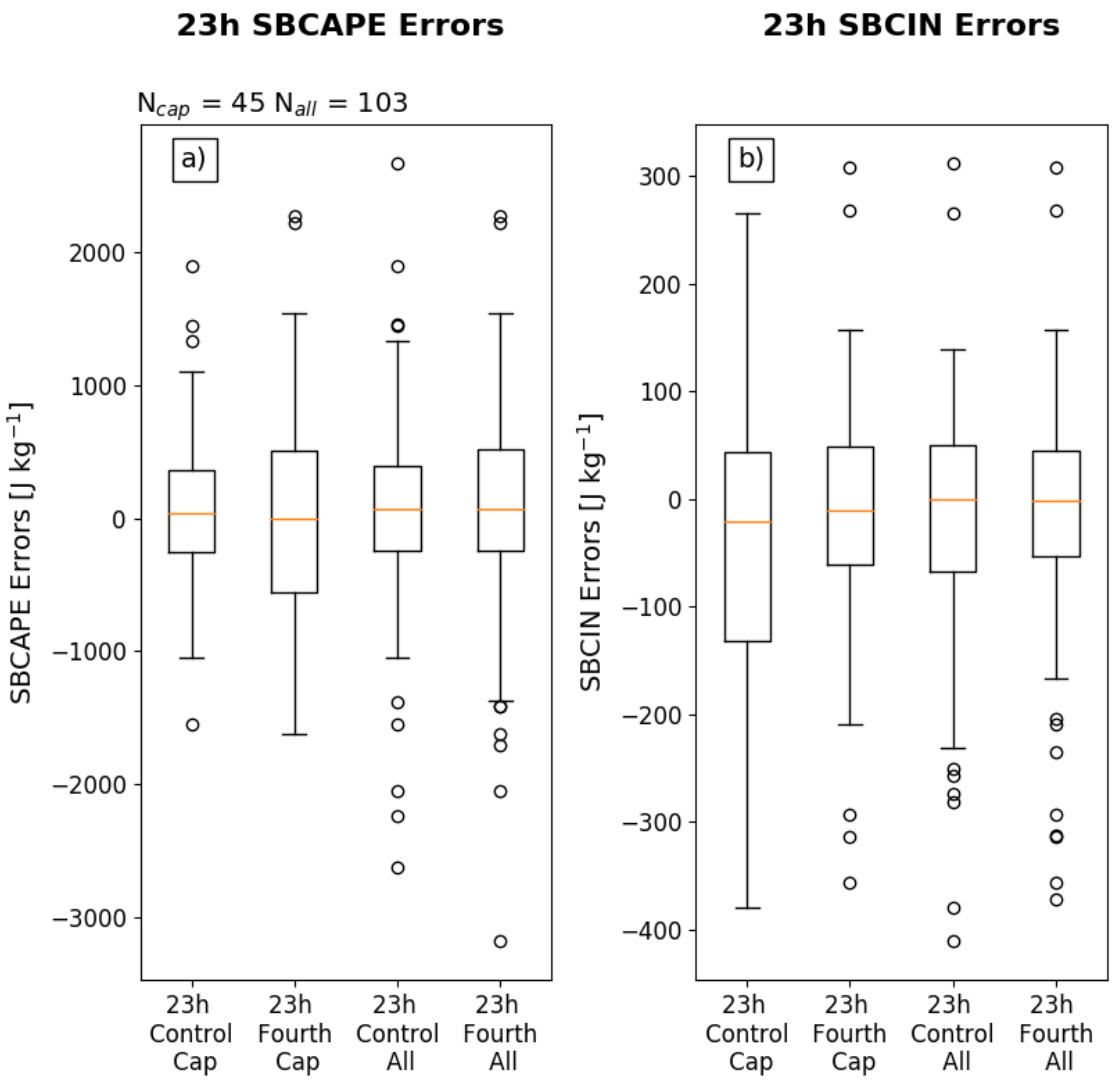

Figure 13. *As in Fig. 8, but for SBCAPE and SBCIN for the 23-h forecast.*

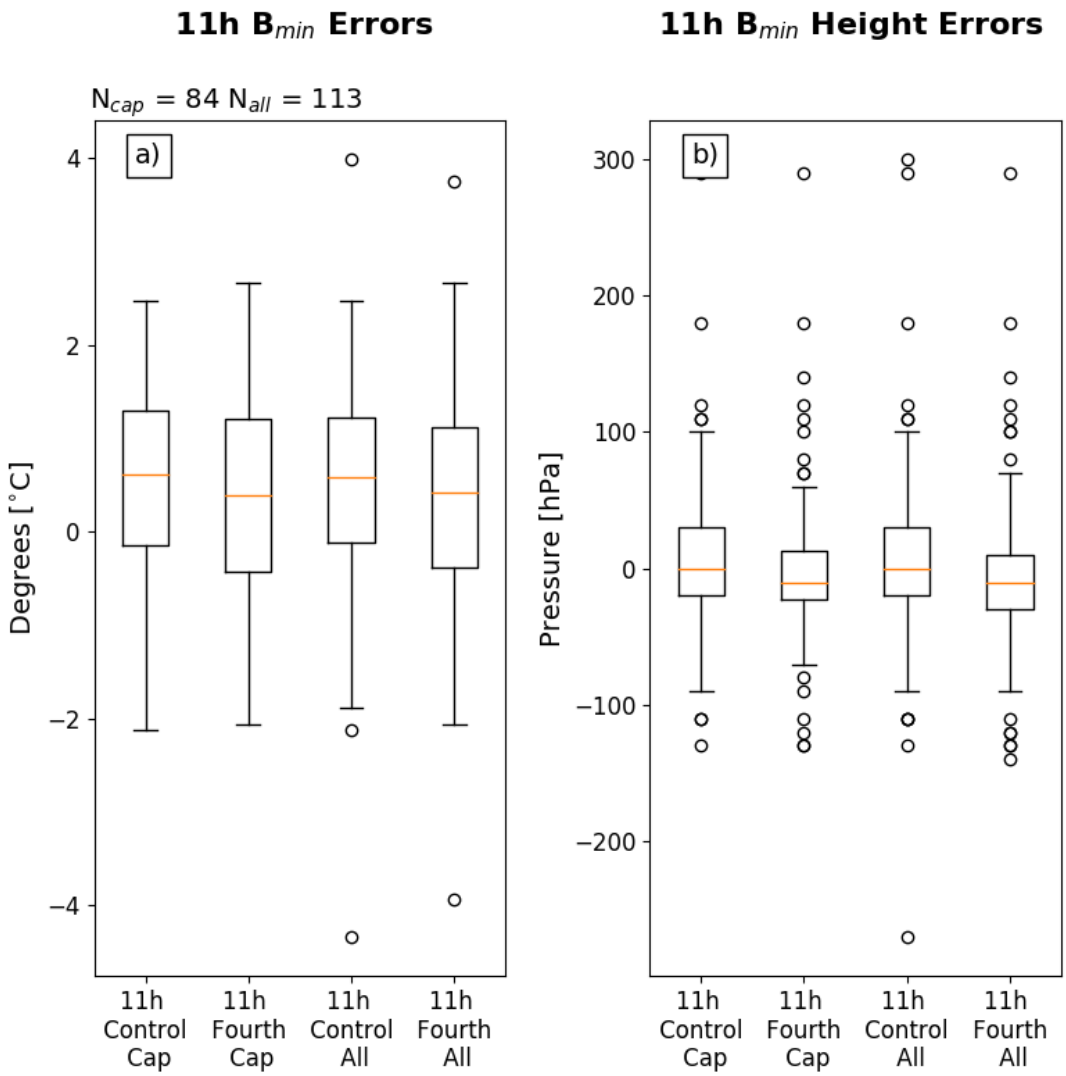

Figure 14. *As in Fig. 8, but for B_{min} and the B_{min} height error for the 11-h forecast. Note that it is still possible to have negative buoyancy but not meet the criteria for a capping inversion.*

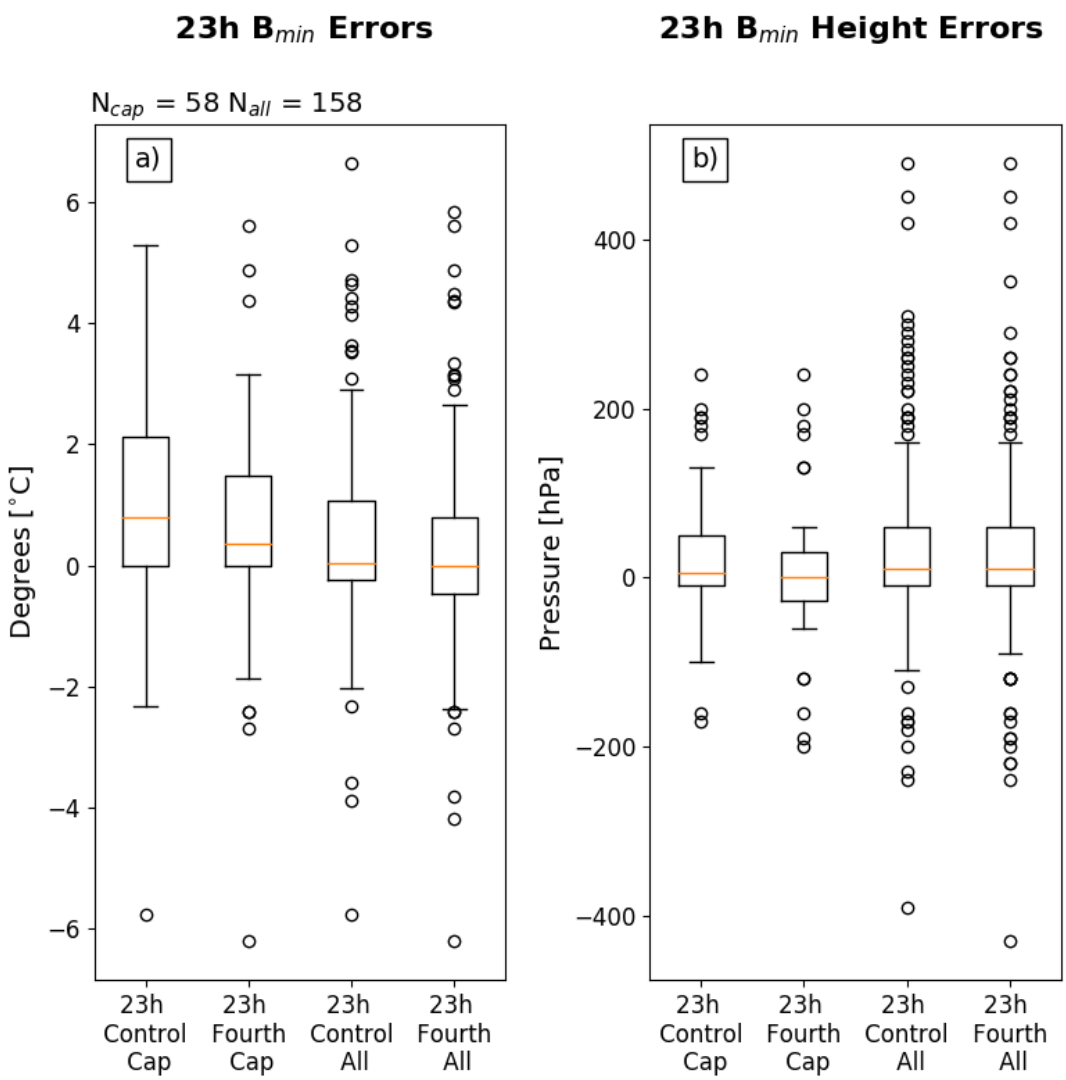

Figure 15. *As in Fig. 14, but for the 23-h forecast.*

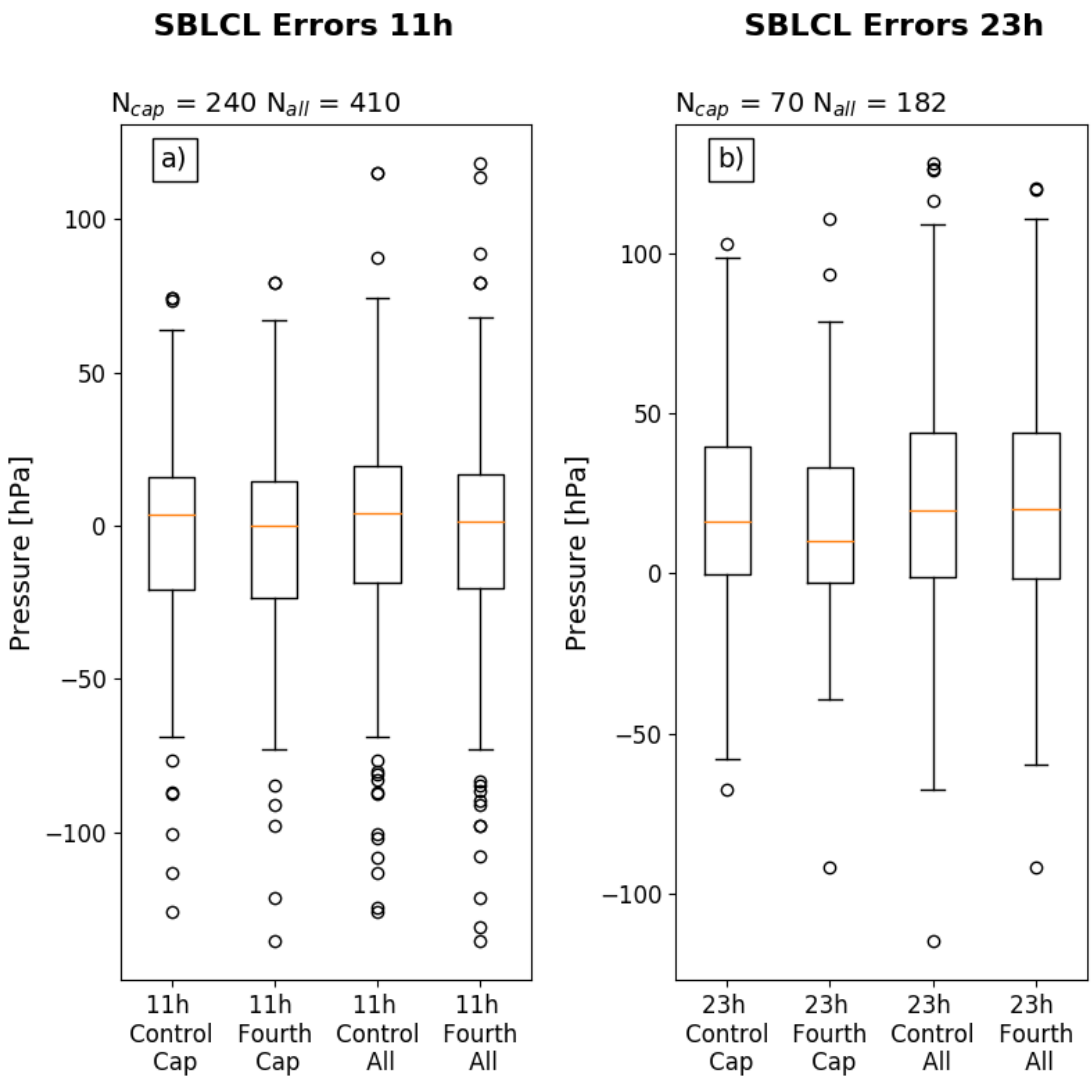

Figure 16. *As in Fig. 8, but for surface-based LCL (SBLCL) for the 11-h (left) and 23-h (right) forecasts.*

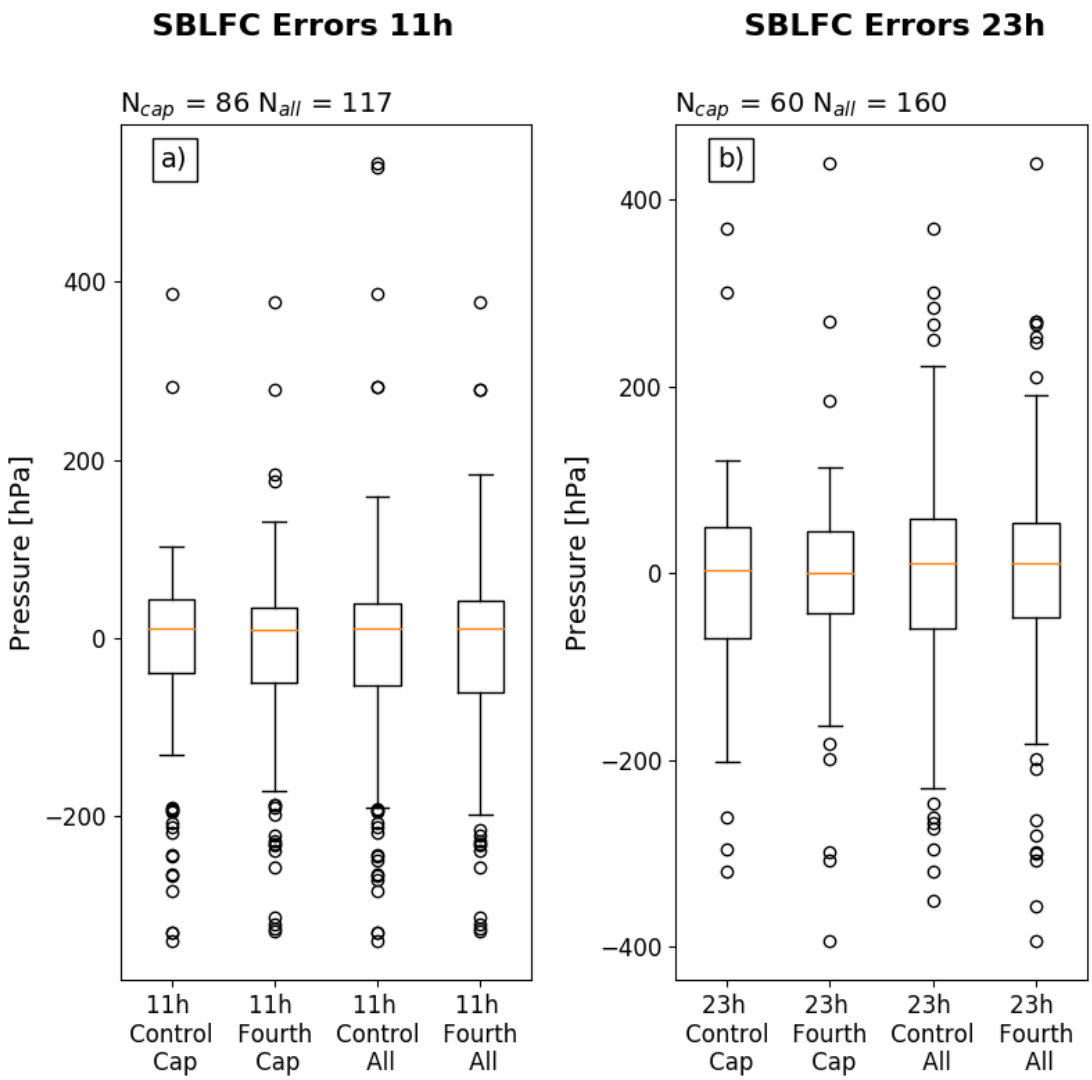

Figure 17. *As in Fig. 8, but for surface-based LFC (SBLFC) for the 11-h (left) and 23-h (right) forecasts.*

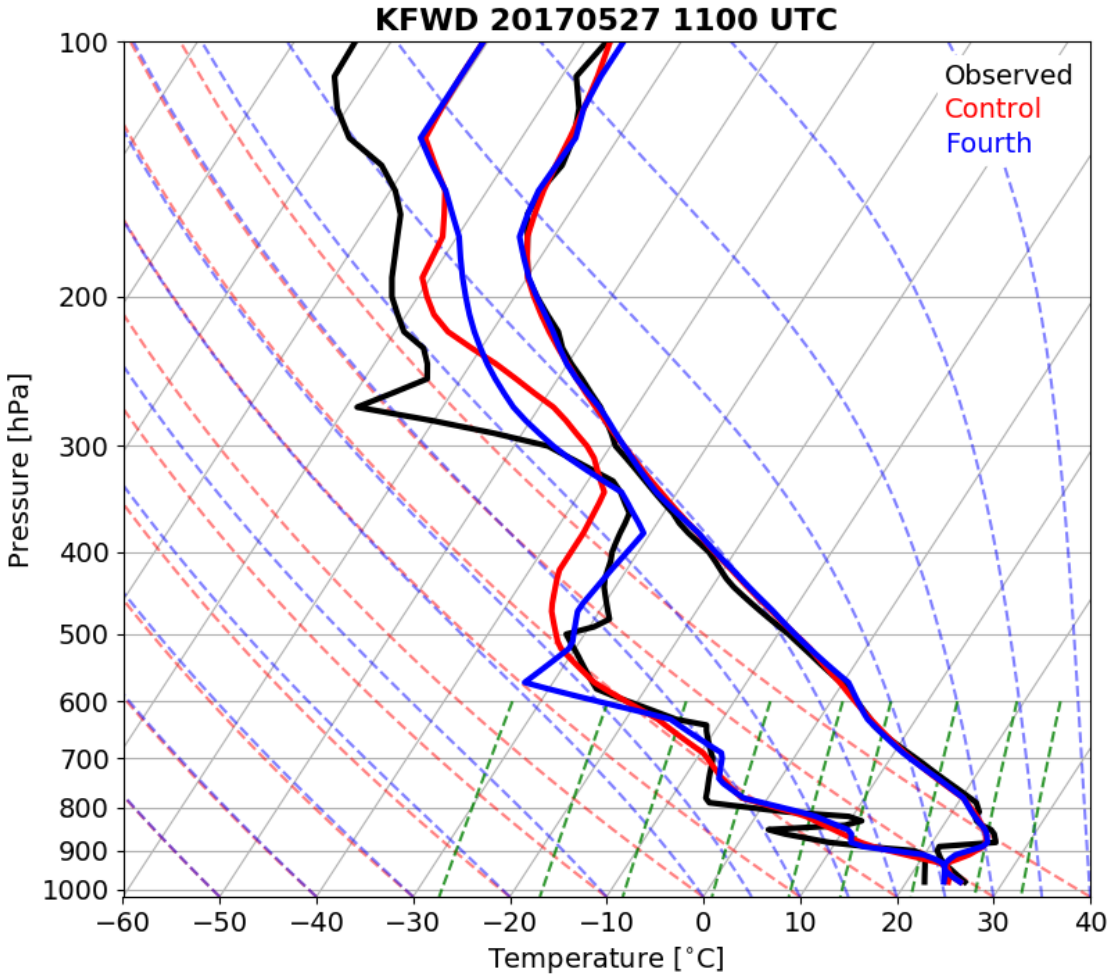

Figure 18. *Skew T–ln p diagram for Dallas-Fort Worth, Texas (KFWD) valid at 1100 UTC 27 May 2017, showing an example of the control and fourth-order simulations artificially smoothing a capping inversion. The red line represents the control simulation sounding, the blue line represents the fourth-order simulation, and the black line is the observed sounding. The left and right traces depict dew point temperature and temperature, respectively. See Table 2 for each profile's derived thermodynamic variables.*

7. References

Benjamin, S. G., and coauthors, 2016: A North American hourly assimilation and model forecast cycle: the Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

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. https://doi.org/10.1175/WAF-D-16-0174.1.

Clark, A. J., and coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74. https://doi.org/10.1175/BAMS-D-11-00040.1.

Clark, A. J., M. C. Coniglio, B. E. Coffer, G. Thompson, M. Xue, and F. Kong, 2015: Sensitivity of 24-h Forecast Dryline Position and Structure to Boundary Layer Parameterizations in Convection-Allowing WRF Model Simulations. *Wea. Forecasting*, **30**, 613–638, https://doi.org/10.1175/WAF-D-14-00078.1

Coffer, B. E., L. C. Maudlin, P. G. Veals, and A. J. Clark, 2013: Dryline position errors in experimental convection-allowing NSSL-WRF forecasts and the operational NAM. *Wea. Forecasting*, **28**, 746–761. https://doi.org/10.1175/WAF-D-12-00092.1.

Cohen, A. E., S. M. Cavallo, M. C. Coniglio, and H. E. Brooks, 2015: A review of planetary boundary layer parameterization schemes and their sensitivity in simulating southeastern U. S. cold season severe weather environments. *Wea. Forecasting*, **30**, 591–612, https://doi.org/10.1175/WAF-D-14-00105.1.

Cohen, A. E., S. M. Cavallo, M. C. Coniglio, H. E. Brooks, and I. L. Jirak, 2017: Evaluation of multiple planetary boundary layer parameterization schemes in southeast U. S. cold season severe thunderstorm environments. *Wea. Forecasting*, **32**, 1857–1884, https://doi.org/10.1175/WAF-D-16-0193.1.

Coniglio, M. C., J. Correia, Jr., P. T. Marsh, and F. Kong, 2013: Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. *Wea. Forecasting*, **28**, 842-862. https://doi.org/10.1175/WAF-D-12-00103.1.

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

Farrell, R. J. and T. N. Carlson, 1989: Evidence for the Role of the Lid and Underunning in an Outbreak of Tornadic Thunderstorms. *Mon. Wea. Rev*., **117**, 857–871, https://doi.org/10.1175/1520-0493(1989)117<0857:EFTROT>2.0.CO;2.

Gallo, B. T., and coauthors, 2017: Breaking new ground in severe weather prediction: the 2015 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Wea. Forecasting*, **32**, 1541–1568. https://doi.org/10.1175/WAF-D-16-0178.1.

Graziano, T. M. and T. N. Carlson, 1987: A Statistical Evaluation of Lid Strength on Deep Convection. *Wea. Forecasting*, **2**, 127–139, https://doi.org/10.1175/1520-0434(1987)002<0127:ASEOLS>2.0.CO;2.

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

Janjic, Z. I., 1994: The step–mountain Eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945. https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Jirak, I., and coauthors, 2015: Spring Forecasting Experiment 2015: preliminary findings and results. [Available at http://hwt.nssl.noaa.gov/Spring_2015/HWT_SFE_2015_Prelim_Findings_Final.pdf.]

Kain, J. S., and coauthors, 2017: Collaborative Efforts between the United States and United Kingdom to Advance Prediction of High-Impact Weather. *Bull. Amer. Meteor. Soc.*, **98**, 937–948, https://doi.org/10.1175/BAMS-D-15-00199.1.

Kain, J. S., P. R. Janish, S. J. Weiss, M. E. Baldwin, R. S. Schneider, and H. E. Brooks, 2003: Collaboration between forecasters and research scientists at the NSSL and SPC: the Spring Program. *Bull. Amer. Meteor. Soc.*, **84**, 1797–1806. https://doi.org/10.1175/BAMS-84-12-1797.

Ladwig, W., 2017: wrf-python (Version 1.1.1) [Software]. Boulder, Colorado: UCAR/NCAR. doi: 10.5065/D6W094P1.

Lanicci, J. M., and T. T. Warner, 1991: A synoptic climatology of the elevated mixed-layer inversion over the southern Great Plains in spring. Part I: structure, dynamics, and seasonal evolution. *Wea. Forecasting*, **6**, 181–197. https://doi.org/10.1175/1520-0434(1991)006<0181:ASCOTE>2.0.CO;2.

Lin, Y. and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: development and applications. Preprints, 19th Conf. on Hydrology, American Meteorological Society, San Diego, CA, 9-13 January 2005, 1.2.

May, R. M., Arms, S. C., Marsh, P., Bruning, E. and Leeman, J. R., 2017: MetPy: A Python Package for Meteorological Data. Unidata, Accessed 15 March 2018. [Available online at https://github.com/Unidata/MetPy; doi:10.5065/D6WW7G29].

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated–k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682. https://doi.org/10.1029/97JD00237.

Powers, J. G., and coauthors, 2017: The Weather Research and Forecasting model: overview, system efforts, and future directions. *Bull. Amer. Meteor. Soc*., **98**, 1717–1737, https://doi.org/10.1175/BAMS-D-15-00308.1.

Skamarock, W. C., and coauthors, 2008: A description of the Advanced Research WRF version 3. *NCAR Tech. Note NCAR/TN-475+STR*, 113pp.

Skamarock, W. C. and M. L. Weisman, 2009: The impact of positive-definite moisture transport on NWP precipitation forecasts. *Mon. Wea. Rev*., **137**, 488–494, https://doi.org/10.1175/2008MWR2583.1

Tewari, M., F., and coauthors, 2004: Implementation and verification of the unified NOAH land surface model in the WRF model. Abstract, *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, Amer. Meteor. Soc., Seattle, WA, 17.

Trapp, R. J., D. J. Stensrud, M. C. Coniglio, R. S. Schumacher, M.E. Baldwin, S. Waugh, and D. T. Conlee, 2016: Mobile radiosonde deployments during the Mesoscale Predictability Experiment (MPEX): rapid and adaptive sampling of upscale convective feedbacks. *Bull. Amer. Meteor. Soc.*, **97**, 329–336, https://doi.org/10.1175/BAMS-D-14-00258.1.

Trier, S. B., C. A. Davis, D. A. Ahijevych, and K. W. Manning, 2014: Use of the parcel buoyancy minimum (B_{min}) to diagnose simulated thermodynamic destabilization. Part I: methodology and case studies of MCS initiation environments. *Mon. Wea. Rev.*, **142**, 945-966. https://doi.org/10.1175/MWR-D-13-00272.1.

Walters, D. N., and coauthors, 2014: The Met Office Unified Model global atmosphere 4.0 and JULES global land 4.0 configurations. *Geosci. Model Dev.*, **7**, 361–386. https://doi.org/10.5194/gmd-7-361-2014, 2014.

Weisman, M. L., and coauthors, 2015: The Mesoscale Predictability Experiment (MPEX). *Bull. Amer. Meteor. Soc.*, **96**, 2127–2149. https://doi.org/10.1175/BAMS-D-13-00281.1.

Wicker, L. J., and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Mon Wea. Rev.*, **130**, 2088-2097. https://doi.org/10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2.

Wood, N., and coauthors, 2014: An inherently mass-conserving semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quart. J. Roy. Meteor. Soc.*, **140**, 1505–1520. https://doi.org/10.1002/qj.2235.