



Article Combining vLAPS and Nudging Data Assimilation

Brian P. Reen ¹,*⁽¹⁾, Huaqing Cai ²⁽¹⁾, Robert E. Dumais, Jr. ², Yuanfu Xie ^{3,†}, Steve Albers ^{3,4,‡} and John W. Raby ²

- ¹ U.S. Army DEVCOM Army Research Laboratory, Adelphi, MD 20783, USA
- ² U.S. Army DEVCOM Army Research Laboratory, White Sands Missile Range, NM 88002, USA; huaqing.cai.civ@army.mil (H.C.); robert.e.dumais.civ@army.mil (R.E.D.J.); john.w.raby2.civ@army.mil (J.W.R.)
- ³ Earth Systems Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA; xieyf@gbamwf.com (Y.X.); Steven.Albers@colostate.edu (S.A.)
- ⁴ Cooperative Institute for Research in the Atmosphere, Boulder, CO 80305, USA
- * Correspondence: brian.p.reen.civ@army.mil
- + Current Affiliation: Guangdong-Hong Kong-Macau Greater Bay Area Weather Research Center for Monitoring Warning and Forecasting (Shenzhen Institute of Meteorological Innovation), Futian District, Shenzhen 518040, China.
- ‡ Current Affiliation: Spire Global, Boulder, CO 80301, USA.

Abstract: The combination of techniques that incorporate observational data may improve numerical weather prediction forecasts; thus, in this study, the methodology and potential value of one such combination were investigated. A series of experiments on a single case day was used to explore a 3DVAR-based technique (the variational version of the Local Analysis and Prediction System; vLAPS) in combination with Newtonian relaxation (observation and analysis nudging) for simulating moist convection in the Advanced Research version of the Weather Research and Forecasting model. Experiments were carried out with various combinations of vLAPS and nudging for a series of forecast start times. A limited subjective analysis of reflectivity suggested all experiments generally performed similarly in reproducing the overall convective structures. Objective verification indicated that applying vLAPS analyses without nudging performs best during the 0–2 h forecast in terms of placement of moist convection but worst in the 3–5 h forecast and quickly develops the most substantial overforecast bias. The analyses used for analysis nudging, and the results suggest that further research is needed on how to best apply analysis nudging of analyses at these scales.

Keywords: variational assimilation; nudging; Newtonian relaxation; 3DVAR

1. Introduction

A variety of techniques to incorporate observations have been used in numerical weather prediction models to allow observations to improve the model forecasts. One example is 3D variational analysis (3DVAR; e.g., [1,2]), and another is Newtonian relaxation (nudging; e.g., [3,4]).

The 3DVAR technique has the potential to determine an optimal analysis by combining observations with a background field. However, since an optimal analysis requires that the background error covariance and the observation error covariance be perfectly known, in practice, the analysis will not be optimal. Background error covariance may be estimated by using model forecasts from a series of past times to find a climatological value, from multiple model forecasts of the current time (i.e., an ensemble) to find case-dependent values, or from a combination of these two methods (e.g., [5]). However, these techniques are difficult to apply when creating on-demand forecasts for an area where equivalently configured forecasts (e.g., horizontal resolution) have not been run in the past and there are insufficient computational resources to carry out an ensemble matching the horizontal grid



Citation: Reen, B.P.; Cai, H.; Dumais, R.E., Jr.; Xie, Y.; Albers, S.; Raby, J.W. Combining vLAPS and Nudging Data Assimilation. *Atmosphere* **2022**, *13*, 127. https://doi.org/10.3390/ atmos13010127

Academic Editor: Stephan Havemann

Received: 16 November 2021 Accepted: 7 January 2022 Published: 13 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). spacing of the on-demand forecasts (although coarser-resolution global ensembles can be used to estimate the case-dependent background error covariance; [5]).

A 3DVAR analysis is commonly applied at a single time at the beginning of a model simulation and can lack physical consistency. Applying the analysis at a single time could result in noise as the model adjusts to a solution consistent with its equations. Additionally, model fields not part of the 3DVAR analyses may not be consistent with the fields in the 3DVAR analyses and may pull the model solution away from the 3DVAR analyses. These potential issues might be mitigated by applying the increments indicated by a 3DVAR analysis over multiple model time steps using incremental analysis updating (e.g., [6,7]) or by nudging towards the 3DVAR analysis. When applying a single analysis, observations made at times other than the analysis time cannot easily be applied at their valid time. While there are methods to mitigate this issue (e.g., first guess at appropriate time, [8]), it is a challenge to know the actual state of the model at the valid time of the observation and to apply the observation at its valid time.

Newtonian relaxation [9,10], also known as nudging, adds nonphysical terms to the tendency terms of the model over a period of time to gradually nudge the model toward analyses or observations. Since the nonphysical terms added by nudging should be smaller than the physical terms in the governing equations, nudging can adjust the model's state variables while maintaining an improved physical consistency compared to inserting modified fields directly at a single time. Nudging allows observations to be applied over a time period centered on their individual valid times, and assimilating analyses valid at multiple times allows observations to be applied to the analysis closest to the valid time of the observation. Fields that are nudged towards must be prognostic variables of the numerical weather prediction model, and thus observations of non-prognostic fields must be converted to prognostic fields in order to be used with nudging. Additionally, nudging is not usually applied in a way that fully accounts for case-specific background error covariances.

Several techniques have combined nudging with another technique. Shaw et al. [11] initialized WRF with a 3DVAR analysis and then used observation nudging and found that the combination appeared to perform better than either solution individually. Liu et al. [12] nudged towards 3DVAR analyses of radar-derived fields while also applying observation nudging. Lei et al. [13–15] combined nudging and the ensemble Kalman filter (EnKF) and applied it first in simplified models and then in the Advanced Research version of the Weather Research and Forecasting model (WRF-ARW; [16]). Information from the ensemble was used to determine how the influence of the observations was spread and how observations of one variable affect another variable. They found the hybrid technique performed better than EnKF for wind direction and better than observation nudging in temperature and relative humidity. Lei et al. [15] also found that noise levels in the hybrid technique were much lower than the EnKF simulation. Liu et al. [17] reported on a hybrid nudging–ensemble system that uses an ensemble to determine the strength at which observation nudging is applied. Lei and Hacker [18] used a Lorenz model to test observation nudging, EnKF, and the nudging–EnKF hybrid developed by Lei et al. [13–15]. They found that in the Lorenz model, the nudging-EnKF hybrid they tested could not simultaneously outperform both nudging and EnKF.

This study demonstrates the combination of a scheme involving 3DVAR (the variational version of the Local Analysis and Prediction System; vLAPS; [19]) and nudging for a single case day with strong convection. This combination takes advantage of vLAPS analysis with the best fit of the observations and nudging with improvement in the physical consistency among the analysis states. Verification of the WRF reflectivity forecasts suggests that further research is needed to improve analysis nudging of high-temporospatial resolution analyses. The experiments described here were previously described in a report [20]; however, that report does not include the objective verification included here. Nonetheless, since this study reports on a subset of the experiments in Reen et al. [20], there are similarities between the text of that document and this study.

This study is unique in its use of analysis nudging to incorporate vLAPS analyses, and there appears to be little past work on assimilating any 3DVAR analyses using analysis nudging. Liu et al. [12] assimilated 3DVAR analyses using analysis nudging but used solely radar observations as the basis for the 3DVAR analyses compared to the much broader set of observations included in the 3DVAR analyses used in our study. While Shaw et al. [11] used 3DVAR and observation nudging, they did not use analysis nudging to apply the 3DVAR analyses, as is done in this study. Analysis nudging is usually used for model forecasts with a horizontal grid spacing much coarser than the 1 km spacing used here (e.g., [21,22]) since analyses are not usually available to nudge towards at such a high resolution. Huo et al. [23] applied what appears to be analysis nudging to assimilate radarderived fields at a fairly high resolution (3 km grid spacing), but one that is somewhat coarser than the 1 km grid spacing in this study. As the current study is limited to a single case day and evaluation of the radar reflectivity field, it is intended as a limited demonstration of the combination of a 3DVAR-based technique (vLAPS) and nudging. One motivating factor in this investigation was to explore assimilation techniques that may be able to provide value for forward-deployed on-demand nowcasting using limited computational capability.

2. Materials and Methods

2.1. WRF-ARW

The Advanced Research version of the Weather Research and Forecasting model (WRF-ARW) V3.6.1 [16] was used with 56 vertical layers over a 1 km-spaced 801 × 801 horizontal grid centered over Oklahoma (Figure 1). The Mellor–Yamada–Janjić scheme (MYJ; [24]) was used to parameterize the atmospheric boundary layer (ABL) with the background turbulent kinetic energy and atmospheric boundary layer depth calculation altered as in Lee et al. [25] and Reen et al. [26]. The Thompson microphysics parameterization [27], the Rapid Radiative Transfer Model longwave scheme [28], the Dudhia shortwave scheme [29], and the Noah land surface model [30] were used.



Figure 1. Areal extent of the Advanced Research version of the Weather Research and Forecasting model (WRF-ARW) 1 km domain. Adapted from Reen et al. [20].

2.2. vLAPS

LAPS [31] uses a modified Barnes analysis [32] with some 1D variational components [33], while vLAPS changes to a multiscale 3DVAR scheme for specific fields (temperature, pressure, winds, and humidity) [19]. The vLAPS 3DVAR scheme is based on the Space and Time Multiscale Analysis System [34] and is multiscale spatially and temporally. The analysis was performed here four times with the coarsest analysis at 16 km horizontally and 50 hPa vertically, and subsequent analyses divided the resolution in half through to the final analysis, which was 2 km horizontally and 25 hPa vertically. The first two analyses used a 30 min time window, whereas the final two analyses used a 15 min time window. The multiscale technique allows observations to spread broadly in data-sparse areas while retaining fine-scale features in more data-rich areas and makes the technique less dependent on the accuracy of the background error covariance. For wind, the control variables were the u- and v-wind components [35] since Xie et al. [34,36] demonstrated that the alternative of the stream function and velocity potential introduce numerical errors and noise. The cloud analysis [33] combined Geostationary Operational Environmental Satellite (GOES) infrared and visible data, radar data, surface observations, and model first-guess fields to construct a cloud fraction. The cloud fraction was used to determine hydrometeor fields and vertical velocity, with the latter affecting the 3D wind. Additionally, the cloud analysis was used as a constraint in determining the vLAPS relative humidity field.

Analyses were created every 15 min using vLAPS. The background field for these analyses was taken from the 3 km-horizontal grid spacing High-Resolution Rapid Refresh (HRRR; [37]); specifically, we used the output from the HRRR 15 UTC cycle on 20 May 2013 (i.e., the integration of HRRR that starts at 15 UTC and provides forecasts at a 15 min temporal spacing). Only the 15 UTC cycle of HRRR was used here to approximate conditions in the application driving this research, namely, use in regions where cycles of the numerical weather prediction model used for initial conditions are available much less frequently than hourly. Each vLAPS analysis used the forecast hour of the 15 UTC HRRR cycle corresponding to the vLAPS analysis time (e.g., the 18 UTC vLAPS analysis used the 3 h forecast of the 15 UTC HRRR cycle). Sources of observations used to create the analyses included Meteorological Assimilation Data Ingest System (MADIS; https://madis.noaa.gov) surface, mesonet, profiler, radiosonde, and Aircraft Communications Addressing and Reporting System (ACARS) observations; pilot observations; WSR-88D radars; and GOES. In addition to the quality control applied by the MADIS data provider, MADIS data were also quality controlled by comparing them to the HRRR fields used as the background.

2.3. Nudging

Nudging [9,10,38] adds a term to the model's tendency equations based on the difference between the current model value and the value from an analysis (analysis nudging; e.g., [39]) or an observation (observation nudging; [40]). This difference is known as the innovation and is multiplied by weighting factors to create the nonphysical addition to the tendency equation. In this study, some experiments applied analysis nudging using vLAPS analysis, and some experiments applied observation nudging of MADIS observations. For brevity, the details of nudging are simplified in the following discussion (see Reen et al. [20] for additional details).

The vLAPS analyses were used by analysis nudging to modify the potential temperature, water vapor mixing ratio, and u- and v-wind components. Within the ABL (atmospheric boundary layer), only the surface analysis was used to calculate analysis nudging terms; at the first level above the ABL, both the surface and above-surface analyses contributed; and above this, only the above-surface analyses contributed. The innovation calculated at the surface (by comparing the vLAPS surface analysis to the WRF surface values) was applied throughout the ABL for the potential temperature and wind. For water vapor, the innovation for each level in the ABL was calculated by comparing the surface analysis to the model value at that level. Innovations above the ABL were calculated by comparing the above-surface analysis to the model value at that level. WRF-ARW analysis nudging linearly interpolates between analyses, which, in this case, were vLAPS analyses available every 15 min. The assimilation period was 3 h, and thus vLAPS analyses from 13 different times were applied. The analysis valid at the end of the assimilation period was applied with a linearly decreasing weight over the 15 min following the end of the assimilation period. The released version of WRF-ARW V3.6.1 does not properly execute this rampdown period, and thus we modified the code to fix this (the fix was provided to the WRF-ARW maintainers and is included in WRF-ARW starting with V3.8). The strength at which analysis nudging was applied was 3×10^{-4} s⁻¹. Analysis (and observation) nudging is designed to gradually modify the model solution while allowing the physical tendency terms to dominate the equations so that the physical consistency of the atmosphere is maintained; this should mitigate any potential issues with overfitting.

Observation nudging used MADIS surface, mesonet, profiler, radiosonde, and ACARS observations (ACARS and profiler observations were inadvertently omitted for part of the simulation). In addition to using the quality control carried out by the data provider, to account for the data quality issues that may be more prevalent in mesonet datasets (e.g., siting issues), mesonet observations were filtered using use/reject lists designed for the Real-Time Mesoscale Analysis [41]. The Obsgrid program [42] was also used to apply quality control. It compared the observations against the 15 UTC cycle of the 3 km HRRR model on 20 May 2013 (i.e., the HRRR forecast that creates forecasts with valid times every 15 min starting at 15 UTC) and against nearby observations. A non-standard version of the WRF Preprocessing System (WPS) program Ungrib was used to vertically interpolate the HRRR field to additional levels to facilitate quality control of single-level above-surface observations (ACARS), and to allow more vertical structure to be retained in multilevel observations (the capability is available in standard Ungrib starting with WPS V3.9).

As with analysis nudging, one must specify the strength of the observation nudging $(6 \times 10^{-4} \text{ s}^{-1} \text{ was used here})$, but additional specification is needed in regard to the vertical, horizontal, and temporal weighting since, unlike the analysis, one does not have one value per model grid cell to nudge towards. Vertical weighting depends on observation type, with the innovation from surface observations being spread throughout the ABL during convective conditions, innovations from single-level above-surface observations being applied in a 100 hPa range, and innovations from multilevel observations being interpolated to model layers within the vertical range of the observation. A 30 km radius of influence was used for surface observations, but terrain differences limited the spreading further. For above-surface observations, the radius of influence increased from 60 km near the surface to 120 km at 500 hPa. Surface observations were used for a 2 h time period centered on the observation valid time, and above-surface observations were used for 3 h time periods. At the end of the assimilation period, observation nudging was ramped down to zero in 1 h (no observations valid during the rampdown were assimilated). The modification of water vapor observation nudging to prevent excessive drying described in [43] was applied. When observation nudging spreads the innovation from a location where the model is too moist to a nearby location that is much drier, this can result in unrealistic drying of the location the innovation is spread to. To mitigate this issue, the modification limits the magnitude of negative water vapor innovations being applied in locations drier than the location where the innovation is calculated.

2.4. Case Description

The day investigated here (20 May 2013) had strong convection in the southern Great Plains region. The base reflectivity (composited over relevant individual radars) of the observed radar (Figure 2) indicated echoes <35 dBZ at 1200, 1500, and 1800 UTC, with most echoes generally along an approximately southwest–northeast-oriented line in the northwest of the domain at 1200 and 1500 UTC, but moving to the north central portion of the domain by 1800 UTC. However, after 1800 UTC, a strong line of southwest-to-northeast-oriented convection developed ahead of the previous echoes and moved eastward. Multiple

tornadoes were observed this day, with one that was determined to be an EF5 on the ground approximately from 1956 to 2035 UTC starting in Newcastle, Oklahoma, and traveling through Moore, Oklahoma [44–46]. Severe hail was also observed in multiple locations.



Figure 2. Observed composite base reflectivity for 20 May 2013 at (**a**) 1200 UTC, (**b**) 1500 UTC, (**c**) 1800 UTC, (**d**) 1900 UTC, (**e**) 2000 UTC, (**f**) 2100 UTC, (**g**) 2200 UTC, and (**h**) 2300 UTC, and for 21 May 2013 at (**i**) 0000 UTC. The white arrow in panel e shows the approximate location of the Moore tornado at that time. Imagery from the Iowa Environmental Mesonet (http://mesonet.agron.iastate.edu/GIS/ apps/rview/warnings.phtml; accessed on 10 November 2021).

Various modeling studies have been performed on this case. Hanley et al. [47] used the United Kingdom Met Office's Unified Model with 4.4/2.2/0.5/0.2/0.1 km horizontal grid spacings without data assimilation and found convection initiating in the Oklahoma City area at approximately the right time, but the tornado-like vortices in their finest grids occurred approximately 2.5 h later than the Moore tornado. Zhang et al. [48,49] used WRF-ARW to look at predictability in this case. Zhang et al. [48], using 27/9/3/1 km horizontal grid spacings, found that temporal shifting of initial conditions generally temporally shifts convection but, in some cases, does not because lateral boundary conditions control convective initiation. Zhang et al. [49] used a 1 km ensemble with perturbations smaller than the current observational network could resolve and found that the ensemble members all produce a line of storms, but that details of individual storms differed. Snook et al. [50] used 500 m-horizontal grid spacing Advanced Regional Prediction System (ARPS)

simulations with assimilation of radar and surface observations every 5 min using EnKF and found skill in predicting hail.

2.5. Experiment Design

The experiments in Table 1 and Figure 3 were used to investigate the potential value of combining vLAPS (2.2) and nudging (2.3) in WRF-ARW (2.1) for the case day of 20 May 2013 (2.4). The names of the experiments begin with the source of the initial conditions (HRRR or vLAPS), followed by the length of the pre-forecast in hours (0 or 3), and finally an O is added if observation nudging was applied, and an A is added if analysis nudging was applied. The pre-forecast refers to the period at the beginning of the model integration during which the model is assumed to be coming into dynamic balance, and during which observations may be assimilated. Hydrometeors are included in the initial conditions provided by both HRRR and vLAPS. Boundary conditions are based on the output from the 15 UTC HRRR cycle. Assimilation of observations or analyses valid during the pre-forecast time may extend into the beginning of the free forecast, but observations or analyses valid during the free forecast should not be assimilated. However, the time window over which observations are included in vLAPS analyses extends 15 min after the valid time of the analysis, and thus observations during the first 15 min of the free forecast are included via the vLAPS analyses.

Table 1. Experimental design. Groups are HS = hot start, WS = warm start, ON = observation nudging, and AN = analysis nudging (VLAPS3AO contains both analysis and observation nudging but is assigned to group AN).

Name	Initial Condition Source	Pre-Forecast Length (h)	Nudgi		
			Analysis	Obs	Group
HRRR0	HRRR	0	Ν	Ν	WS
HRRR3	HRRR	3	Ν	Ν	WS
VLAPS0	vLAPS	0	Ν	Ν	HS
VLAPS3	vLAPS	3	Ν	Ν	WS
VLAPS3O	vLAPS	3	Ν	Y	ON
VLAPS3A	vLAPS	3	Y	Ν	AN
VLAPS3AO	vLAPS	3	Y	Y	AN

The experiments initialized with HRRR0 and HRRR3 served as the no-assimilation control experiments for the 0 h pre-forecast and 3 h pre-forecast experiments, respectively. HRRR0 and HRRR3 differ from one another in that, for a given cycle, the HRRR3 experiment starts integrating 3 h earlier to allow WRF to spin up. For example, the 20 UTC cycle of HRRR0 (referred to as HRRR0₂₀) starts integrating at 20 UTC using the 5 h forecast from the 15 UTC HRRR cycle, whereas HRRR3₂₀ starts integrating at 17 UTC using the 2 h forecast from the 15 UTC HRRR cycle, but the output between 17 and 20 UTC is not considered part of the forecast since it is during the spinup time. The experiments initialized with vLAPS that applied no additional nudging were used to determine the potential value of adding nudging. VLAPS0 represents the normal way in which vLAPS (and other 3DVAR analyses) is used, while VLAPS3 allows the effect of a 3 h pre-forecast without additional nudging assimilation to be ascertained. Both experiments used the vLAPS analysis as the initial conditions for the model integration. The other three experiments used vLAPS for the initial conditions, and during a 3 h pre-forecast, they analysis nudged toward the thirteen vLAPS analyses valid during this period (VLAPS3A), observation nudged toward observations (VLAPS3O), or both (VLAPS3AO). These experiments explore the value of applying a series of vLAPS 3D analyses rather than just one, and whether observation nudging might add value even when vLAPS analyses are used.



Figure 3. Experimental design. The initial conditions for the experiment are labeled in white letters. The short vertical blue lines represent the valid times of the variational version of the Local Analysis and Prediction System (vLAPS) analyses which were used in analysis nudging for a time period centered on the valid time. The rampdown period of nudging as described in Section 2.3 is omitted from this figure for clarity. Each experiment was run with t_0 values of 1800, 1900, 2000, 2100, 2200, and 2300 UTC. Adapted from Reen et al. [20].

For the purpose of comparing the experiments in the objective verification section, we broke them up into groups based upon the general method of data assimilation applied. These groups are referred to as hot start (HS; VLAPS0), analysis nudge (AN; VLAPS3A and VLAPS3AO), observation nudge (ON; VLAPS3O), and warm start (WS; HRRR0, HRRR3, and VLAPS3). The experiment that used both analysis and observation nudging was placed within the AN group, and the WS group included all experiments that "spun up" from either a vLAPS- or HRRR-generated background state without further data assimilation application. The motivation for including the non-vLAPS experiments in the WS group was that the HRRR 15 UTC forecast cycle fields provide sufficiently "spun up" mesoscale information (including hydrometeor fields), due to the model's own high-quality data assimilation methodology and 3 km native grid spacing, meaning it is essentially a warm start. In summary (Table 1), HS included VLAPS0, AN included VLAPS3A and VLAPS3AO, ON included VLAPS3O, and WS included HRRR0, HRRR3, and VLAPS3.

For each experiment, independent cycles were carried out hourly with the integration period of all cycles ending at 0000 UTC on 21 May 2013 (Figure 4). The 0 h forecast time is the end of the pre-forecast and the beginning of the forecast (t_0 in Figures 3 and 4), which

was 1800 UTC for the first cycle and 2300 UTC for the final cycle (i.e., cycles were run with the following t_0 values: 1800, 1900, 2000, 2100, 2200, and 2300 UTC). Each cycle is referred to by its 0 h forecast time (t_0). Each WRF cycle includes the integration of WRF through the 3 h pre-forecast period up to t_0 (except for VLAPS0 and HRRR0, which do not have a pre-forecast period) and the integration of WRF from t_0 until the end of the forecast at 0000 UTC. For example, the 18 UTC cycle of VLAPS3AO (referred to as VLAPS3AO₁₈) starts integration at 1500 UTC, applies analysis and observation nudging data assimilation during the integration in the 3 h pre-forecast period ending at 1800 UTC (t_0), and then continues integrating until the end of the forecast period at 0000 UTC. Each cycle is independent in that it is not affected by any of the other cycles.



Figure 4. Illustration of the six WRF cycles used for each experiment and which are referred to by the 0 h forecast time (t_0). As shown in Figure 3, the 3 h pre-forecast period wherein any analysis or observation nudging data assimilation was applied is omitted in some experiments (HRRR0 and VLAPS0). The hourly forecast lead times for each cycle are labeled in red; the output was created at a 15 min temporal spacing.

All experiments used the 1500 UTC cycle of HRRR as the starting point to determine the initial conditions and boundary conditions, with the vLAPS analyses using this cycle as the first-guess field. This simplifies the experimental design and is more representa-

tive of the limited frequency of updated model data available to drive finer-scale model simulations in battlefield conditions.

In order to more concretely illustrate the experimental configuration, Figure 5 shows how one specific experiment (VLAPS3AO) was configured for one specific cycle (2100 UTC), i.e., VLAPS3AO₂₁. The lower portion of Figure 5 illustrates how analysis nudging was applied during the 3 h pre-forecast (1800–2100 UTC for this cycle) using 15 min vLAPS analyses by showing the analysis nudging details for the first hour of this period (i.e., 1800–1900 UTC). Each analysis was applied to the model over 30 min centered on the valid time of the analysis. The weight at which it was applied linearly increased in the 15 min prior to the valid time of the analysis and linearly decreased in the 15 min following the valid time. Observation nudging applied observations over a time window centered on each individual observation; as described in Section 2.3, a 2 h time window was used for surface observations.



1800 - 1900 UTC Magnified to Demonstrate Analysis Nudging Configuration

Figure 5. Schematic illustrating the details of a specific experiment (VLAPS3AO) for a specific cycle (21 UTC), i.e., VLAPS3AO₂₁. While the lower half of the figure shows the details of how vLAPS analyses were applied using analysis nudging between 1800 and 1900 UTC, analysis nudging towards vLAPS analyses took place throughout the 3 h pre-forecast period.

3. Results

Here, the evaluation of the combination of a 3DVAR-based technique (vLAPS) and nudging is limited to reflectivity. To demonstrate the variation among the experiments, a single forecast time for a single cycle is examined subjectively. However, given the number of experiments, cycles, and forecast times, the objective evaluation that follows allows for a more holistic understanding of the results.

3.1. Subjective Evaluation

The observed composite base reflectivity is compared to the WRF-ARW lowest-level reflectivity for the 18 UTC cycle of each experiment in Figure 6 for 2015 UTC 20 May 2013. This is near the time of the peak of the EF5 tornado and was chosen as a time where strong convection is present. The observed reflectivity (Figure 6c) shows convection along a southwest-to-northeast-oriented line from Texas through the southeast corner of Kansas. There are weaker echoes (<30 dBZ) in a short line in central Kansas, and scattered echoes in the northwest corner of the domain. All of the experiments forecast the main line of convection and some echoes in the northwest corner of the domain, but there are notable differences among the experiments.



Figure 6. Reflectivity at 2015 UTC on 20 May 2013 from (**a**) HRRR0₁₈; (**b**) VLAPS0₁₈; (**c**) observations; (**d**) HRRR3₁₈; (**e**) VLAPS3₁₈; (**f**) VLAPS3A₁₈; (**g**) VLAPS3O₁₈; (**h**) VLAPS3AO₁₈. The black X within the white square shows the approximate location of the EF5 tornado. The observed reflectivity is the composite base reflectivity, and the WRF-ARW reflectivity is the lowest model-level radar reflectivity. The observed reflectivity is from the Iowa Environmental Mesonet (http://mesonet.agron.iastate.edu/GIS/apps/rview/warnings.phtml; accessed on 10 November 2021).

The northern portion of the main line of convection is not well simulated by the experiments initialized with HRRR (HRRR0₁₈ in Figure 6a and HRRR3₁₈ in Figure 6d). The use of vLAPS as the initial condition slightly strengthens convection in this area (VLAPSO₁₈ in Figure 6b), and the addition of the 3 h pre-forecast without nudging (VLAPS3₁₈ in Figure 6e) brings the strength closer to the observations. However, it appears that the experiment analysis nudging to the vLAPS analyses and applying observation nudging (VLAPS3AO₁₈ in Figure 6h) best reproduces the continuous strength of the convection along this line in Kansas. The southern edge of the main line of convection does not extend far enough south in VLAPS3₁₈ (Figure 6e), but most of the other experiments appear to extend it too far south. All experiments show echoes in the general region of the observed EF5 tornado, but only three experiments (VLAPSO₁₈ in Figure 6b, VLAPS3₁₈ in Figure 6e) show echoes at the location, with two of these experiments (VLAPSO₁₈ in Figure 6b and VLAPS3O₁₈ in Figure 6g) showing echoes ≥ 40 dBZ very close to the location.

The area of weak echoes in central Kansas is not well forecast by any of the experiments. The HRRR-initialized experiment without a pre-forecast (HRRR0₁₈ in Figure 6a) shows the feature stronger and smaller than observed and, unlike the observations, does not show an area with no echoes between it and the echoes in the northwestern portion of the domain. The vLAPS-initialized experiment without a pre-forecast (VLAPS0₁₈ in Figure 6b) shows a pair of lines of convection in this area that are both much stronger than observed. The 3 h pre-forecast (HRRR3₁₈ in Figure 6d) allows the HRRR initialization to reproduce more of the separation between the feature and echoes in the northwestern portion of the domain observed but shows the feature with much less coverage than observed. For experiments with the 3 h pre-forecast, using vLAPS as the initial conditions broadens the area covered by this feature compared to the HRRR initialization (HRRR3₁₈) and removes the second line of convection seen in the vLAPS experiment without the 3 h pre-forecast (VLAPSO₁₈). The experiments with the area covered by this feature that most closely match the observations appear to be the experiments analysis nudging the vLAPS analyses (VLAPS3A₁₈ in Figure 6f and VLAPS3AO₁₈ in Figure 6h). However, all of the experiments that produce the feature show it stronger than observed. The echoes in the northwest quadrant of the domain, in general, appear to be stronger than observed in the experiments, perhaps most so for experiments analysis nudging towards vLAPS analyses or observation nudging (VLAPS3A₁₈ in Figure 6f, VLAPS3O₁₈ in Figure 6g, and VLAPS3AO₁₈ in Figure 6h).

All of the experiments produce model base reflectivity fields that are at least generally consistent with the observed composite radar reflectivity, but there is notable variation among the experiments. While it is difficult to subjectively determine which experiment performs best at the time shown here for this cycle, the results suggest that the experiment analysis nudging to the vLAPS analyses and applying observation nudging performed reasonably well (VLAPS3AO₁₈ in Figure 6h).

3.2. Objective Evaluation

The previous subsection presented subjective support that the model simulations realistically reproduced the general convective outbreak event on the afternoon of 20 May, including aspects associated with the timing, spatial location, alignment, and convective mode. Reen et al. [20] showed that the model had some ability in capturing realistic gross structures seen in supercellular storms. However, the evolution of the convective outbreak and structural details of individual convective elements differed among the experiments; this variation offers a spread of potential solutions, consistent with what might be expected based on the findings in Zhang et al. [49]. The following section provides a more objective evaluation using several forecast performance measures to compare all of the different experiments aggregated across the hourly cycles during the afternoon of May 20.

Although the results are only valid for this single unique case study event, they may provide valuable clues as to how the various techniques can perform in a more general sense. We are also interested in whether our combined 3DVAR analysis/observation nudging strategy might be of value in supporting stand-alone operational nowcasting systems run on a more modest computational hardware platform.

3.2.1. Metrics Used in Objective Evaluation

To quantitatively measure the performances of each cycle across the full spectrum of our experiments, a trio of well-established forecast evaluation measures was used to examine the short-range convective forecasts by comparing observed to model forecast radar reflectivity fields. The fractions skill score, or FSS [51], is a spatial verification metric often used for assessing the performance of precipitation forecasts from numerical weather prediction (NWP) models (frequently for convective precipitation). The critical success index, or CSI [52], is another common metric used to evaluate categorical forecasts, taking into account hits, misses, and false alarms. Finally, the frequency bias score (FBIAS) [53,54] is also used to assess the quality of model-derived radar reflectivity field forecasts against real radar observations [19,53]. The FBIAS is simply the ratio between the number of model grid points where the model forecasts reflectivity above a certain threshold and the number of model grid points where the observed radar reflectivity values exceed the same threshold.

The FSS was computed for different neighborhood sizes aggregated from the native grid resolution. The aggregation was also conducted across all of the relevant cycles for each lead forecast time in each experiment (not all cycles are relevant for each forecast lead time since some cycles are too short to include some forecast lead times). For example, the FSS for the 1.25 h forecast lead time of VLAPS3AO aggregates the verification of VLAPS3AO₁₈ at 1915 UTC, VLAPS3AO₁₉ at 2015 UTC, VLAPS3AO₂₀ at 2115 UTC, and VLAPS3AO₂₁ at 2215 UTC (Figure 4 shows the relationship between forecast lead time and valid time for each cycle). Since model forecast skill often varies based on the forecast lead time, aggregating statistics by forecast lead time reveals how the temporal evolution of error varies among methodologies. The CSI and FBIAS were not computed using neighborhoods but were both calculated on the highest available resolution grid (here, our native model grid spacing of 1 km) and, as with FSS, were aggregated across cycles for each lead forecast time in each experiment. The FBIAS of reflectivity measures how much the model either overforecasts (forecast is more than the actual) or underforecasts (forecast is less than the actual) coverage of reflectivity values exceeding a threshold. No verification takes place during the pre-forecast (where data assimilation occurs), but verification is instead limited to the forecast period.

As noted previously (and shown in Figure 4), a model integration with a t_0 of each hour from 1800 UTC to 2300 UTC was launched (i.e., an hourly refresh) for each experiment and produced 15 min forecasts from t_0 out to 0000 UTC of 21 May. The latest time verified is 2300 UTC because radar data provided by vLAPS were not available after this time. The output was produced in 15 min intervals, which thus produced forecasts for verification with lead times ranging from 15 min to 5 h (resulting in statistics at 20 different forecast lead times). Due to the 0000 UTC end time of all cycles, but radar data not being available after 2300 UTC, only the 1800 UTC cycle produced a 5 h forecast lead time. Because of various pre-forecast requirements across different experiments, some experiments start their integration before their 0 h lead time (e.g., a model cycle with a base time of 1800 UTC with a 3 h pre-forecast requirement was actually initialized at 1500 UTC with lead times of forecasts based on the 1800 UTC mark). Since all simulations terminated at 0000 UTC, the model integrations which started at 2300 UTC had only a maximum 1 h lead forecast available. Because radar data were not available after 2300 UTC, all verification forecast times of the 2300 UTC cycle (other than the 0 h forecast time) are after the end of radar data availability, and thus the 2300 UTC cycle is not included in the objective verification statistics. WRF-ARW "restart" forecasts were not used to initialize from a previous cycle's forecast—all first-guess initial condition (and lateral boundary tendency) fields for all cycles leveraged the 1500 UTC HRRR cycle forecast fields available at hourly increments. A single HRRR cycle was used to more closely approximate the cycle frequency that is available outside of the United States for models such as the Global Forecast System that could be used as boundary conditions for fine-scale simulations.

Each experiment was aggregated across its simulations to create statistics, with each value representing that experiment's performance for the forecast lead time, reflectivity threshold, and (for FSS) neighborhood size being evaluated. Similar to other studies, the column maximum radar reflectivity from both the model and the National Weather Service WSR-88Ds (provided by the National Oceanic and Atmospheric Administration; NOAA) was used to evaluate the convective forecasts [55]. The ground truth column maximum radar reflectivity fields obtained from the NOAA are the same type as those used for previous studies involving the original LAPS product [33]. The model column maximum reflectivity was computed using the WRF-ARW diagnostic method [56] that uses prognostic hydrometeor fields from the model microphysics parameterization.

The FSS for two radar reflectivity thresholds (25 dBZ and 35 dBZ) using a 9 km neighborhood size is shown in Figure 7. The CSI and FBIAS metrics are shown in Figures 8 and 9. The neighborhood size of 9 km is here considered to be a reasonable estimate for the "effective resolution" of the WRF-ARW when run on a mesh with a native 1 km grid spacing [57]. The number of cycles included in the statistics varies by forecast lead time, with five cycles included at the 1 h forecast and one cycle included at the 5 h forecast.

3.2.2. Outcome of Objective Evaluation

For the purpose of comparing the experiments in this section, we utilized the groups described in Section 2.5 which are based upon the general method of data assimilation applied. The various FBIAS, FSS, and CSI curves throughout the lead forecast period of study (as seen in Figures 7–9) show similarity within these assimilation methodology groups, namely, hot start (HS), analysis nudge (AN), observation nudge (ON), and warm start (WS). Note that the time subscript is omitted when referring to the evaluation of the WRF forecasts in this section because each experiment is evaluated based on data that are combined across the cycles.

Although VLAPS0 is considered as a HS, it does differ from previous LAPS/vLAPS HS experiments [19,58] in that it did not directly insert the vertical velocity from the analysis into the WRF-ARW initial conditions. However, the hydrometeor fields produced in the vLAPS analyses are used in the appropriate initial WRF-ARW arrays. Additionally, since both reflectivity and radial wind information from the WSR-88D radars were incorporated into the vLAPS analyses, it is expected that the initial vertical velocity fields adjusted quickly through continuity considerations to the convective-scale information provided by the radar data. Any experiment that used a pre-forecast period and which started from an initial vLAPS analysis also filled the initial WRF-ARW arrays in the same fashion (Section 2 provides more details on the use of vLAPS analyses for model initialization). When the vLAPS analyses are used instead solely as a source for 15 min intermittent analysis nudging across a pre-forecast period by WRF-ARW, the analyses are only used to nudge the potential temperature, water vapor mixing ratio, and u- and v-wind components, as these are the variables for which analysis nudging is available in WRF-ARW.



Figure 7. FSS by lead forecast time (by experiment) and for (**a**) 25 dBZ and (**b**) 35 dBZ reflectivity threshold levels (for a 9 km neighborhood size). Line colors indicate whether an experiment is in the warm start (WS), hot start (HS), analysis nudge (AN), or observation nudge (ON) group. The number of cycles included in the statistics at each integer lead time hour is shown along the top of the figure.



Figure 8. CSI by lead forecast time (by experiment) and for (a) 25 dBZ and (b) 35 dBZ reflectivity threshold levels. Line colors indicate whether an experiment is in the warm start (WS), hot start (HS), analysis nudge (AN), or observation nudge (ON) group. The number of cycles included in the statistics at each integer lead time hour is shown along the top of the figure.



Figure 9. FBIAS by lead forecast time (by experiment) and for (**a**) 25 dBZ and (**b**) 35 dBZ reflectivity threshold levels. Line colors indicate whether an experiment is in the warm start (WS), hot start (HS), analysis nudge (AN), or observation nudge (ON) group. The number of cycles included in the statistics at each integer lead time hour is shown along the top of the figure.

Upon examining Figures 7–9, the first thing of immediate note is that for this case study event, the HS experiment stands out from the others, mostly within the initial lead forecast hour. The FSS and CSI scores for the HS experiment are clearly superior to all other groups during at least the first 30 min of lead forecast time (independent of whether

the 25 dBZ or 35 dBZ reflectivity threshold value is used). This strong advantage for very short term nowcasting of vLAPS is most likely due to its use of a diabatic hot start methodology and has been noted previously [19,58]. The motivation with vLAPS has always been to compete with and exceed the skill of simpler nowcasting methods such as feature-based advection/extrapolation or basic persistence during the initial hour. This is a very challenging goal for convective-scale NWP modeling. In at least this study, the flip side to HS seems to be that the FSS and CSI scores quickly degrade back to those of the other methods by approximately a 2 h lead time and, by a 5 h lead time, even appear to result in somewhat worse scores. Perhaps this is some evidence of an overfitting to the stronger reflectivity echoes at time 0 h, and that some degree of imbalance could still remain between the full set of mass and momentum model fields produced by the hot start analysis itself.

The AN group of experiments exhibited the highest FSS and CSI scores during the first lead hour (outside of HS). The AN group shows that both FSS and CSI gradually declined to approximately 2 h and then remained mostly flat until approximately 4 h. During most of the period after approximately the 1 h lead forecast time, the AN experiments clump in with the other groups in terms of general FSS and CSI scores. An exception is the VLAPS3AO experiment in the AN group, which shows a gradual decrease in FSS and CSI scores after the 4 h lead forecast—all other experiments in all groups show a slight increase in FSS and CSI scores starting at approximately the 4.5 h lead forecast. This may be an artifact of the very limited number of cycles (and thus very small sample size) available to compute metrics at the 4 h and 5 h lead forecasts. It could also point to some need to fine-tune aspects of the combined analysis and observation nudging criteria used by VLAPS3AO, particularly when combining meso/synoptic-scale conventional observations via observation nudging with convective-scale observation data assimilated through analysis nudging to the 15 min updated vLAPS analyses (when the radar data are incorporated).

The ON group fell in the middle to lower half of the pack of all groups, both for the short and longer lead forecast times, in terms of FSS and CSI scores. The FSS and CSI scores for the VLAPS3O experiment generally remained fairly consistent at all lead forecast times.

The FSS and CSI scores for the WS group tended to track fairly closely to those of the ON group, although HRRR0 was a bit lower initially, while VLAPS3 beyond 2.5 h was a notable exception (higher FSS at further lead times). For the 25 dBZ threshold, the WS group (including HRRR0 and HRRR3) tended to outperform the ON group in FSS and CSI through much of the lead forecast period after approximately 3 h. VLAPS3 generally showed the highest FSS and CSI scores of all experiments from lead forecasts of 2.5 h and beyond, particularly for the 25 dBZ reflectivity threshold.

Tables 2 and 3 show the FSS scores at lead forecast times of 0.25 h, 1.75 h, 3.50 h, and 5.00 h for each experiment, and for the 25 dBZ and 35 dBZ thresholds, respectively (for a 9 km neighborhood size). The tables more clearly show the data values, which can also be estimated from Figure 7.

T	FSS by Lead Forecast (h)				
Experiment —	0.25	1.75	3.50	5.00	
HRRR0	0.22	0.34	0.38	0.42	
HRRR3	0.33	0.34	0.35	0.41	
VLAPS0	0.75	0.40	0.30	0.22	
VLAPS3	0.29	0.33	0.42	0.53	
VLAPS3A	0.44	0.40	0.38	0.41	
VLAPS3AO	0.41	0.39	0.33	0.26	
VLAPS3O	0.28	0.35	0.32	0.46	

Table 2. FSS for 25 dBZ threshold and 9 km neighborhood size, at four different lead forecast times (h).

Evenories ant	FSS by Lead Forecast (h)				
Experiment –	0.25	1.75	3.50	5.00	
HRRR0	0.11	0.22	0.22	0.28	
HRRR3	0.20	0.22	0.21	0.28	
VLAPS0	0.72	0.23	0.16	0.11	
VLAPS3	0.18	0.23	0.30	0.34	
VLAPS3A	0.33	0.25	0.24	0.26	
VLAPS3AO	0.31	0.24	0.20	0.11	
VLAPS3O	0.19	0.28	0.22	0.31	

Table 3. FSS for 35 dBZ threshold and 9 km neighborhood size, at four different lead forecast times (h).

Examining the FBIAS among the groups (Figure 9), we can determine whether the model was forecasting too much (overforecasting; FBIAS > 1.0) or too little (underforecasting; FBIAS < 1.0) areal coverage of reflectivity at or above a given reflectivity threshold. Early in the lead forecast period, nearly all the groups of experiments showed a significant tendency to overforecast the stronger convection (higher reflectivity) areas, such as at the higher 35 dBZ threshold. HRRR0 from the WS group produced the least biased FBIAS results for both thresholds early in the lead forecast time, but with increasing FBIAS through approximately the 3 h lead forecast. The VLAPS3 and HRRR3 experiments, also from the WS group, tended to also be less high biased during the early lead forecast period. All of the other experiments from the different groups started in mostly the 2.0 to 3.0 FBIAS range for both the 25 dBZ and 35 dBZ thresholds but dropped steadily towards FBIAS values of 1.0 to 1.5 by the end of the 5 h lead forecast period. At the 25 dBZ threshold, the ON group began to underpredict areal coverage by the end of the forecast period (VLAPS3O).

At the 25 dBZ and 35 dBZ thresholds, overforecasting was the main forecast feature across all experiments, particularly early in the lead forecast period. The single HS experiment, VLAPS0, showed the most exaggerated overforecasting at approximately the 30 to 45 min lead forecast for the 35 dBZ threshold (as high as approximately 3.25) but dropped steadily towards \approx 1.25 by the end of the 5 h forecast period. The experiments VLAPS3AO, VLAPS3A, and VLAPS3O followed a similar FBIAS evolution (in terms of rate of increase and decrease) to VLAPS0 for both the 25 dBZ and 35 dBZ thresholds after the 1 h lead time. During the 0 h to 1 h lead forecast period, the FBIAS decreased in the VLAPS3O forecasts, whereas the VLAPS3A and VLAPS3AO FBIAS remained fairly constant. In contrast to all of the other experiments, the VLAPS0 FBIAS sharply increased between the 0.25 h and 0.50 h forecasts, suggesting that the initial analysis solution may not be in complete balance (perhaps from overfitting to the radar observations) with the WRF-ARW model on the 1 km domain. Tables 4 and 5 show the specific FBIAS scores at lead forecast times of 0.25 h, 1.75 h, 3.50 h, and 5.00 h for each experiment, and for the 25 dBZ and 35 dBZ thresholds, respectively. The overforecasting at the 25 dBZ and 35 dBZ thresholds is generally more pronounced for forecasts recently influenced by observations (via vLAPS analyses directly inserted at t_0 in VLAPS0, via vLAPS analyses being analysis nudged towards in VLAPS3A and VLAPS3AO, or via observation nudging in VLAPS3O). It may be that the methods used to incorporate observations into vLAPS analyses and through observation nudging make assumptions that do not work optimally for this specific case and the specific observations available; future work should explore the generality of this bias and seek to determine how the methodologies could be improved to lower the overforecast bias.

The variation in performance by neighborhood size and reflectivity threshold (as measured by FSS) for this case study can be seen more clearly in Table 6. In Table 6, the FSS is shown for a few different lead forecast times near the start and end of the forecast period (0.50 h and 5.00 h), and for different neighborhood sizes (1 km, 9 km, 17 km). The lower 10 dBZ threshold is focused upon in Table 6, although for the 9 km neighborhood (closest to the effective resolution of the WRF 1 km grid spacing output), additional thresholds of 25 dBZ and 35 dBZ are also shown. The tendency of FSS to improve with lower reflectivity

thresholds and larger neighborhood sizes, as seen in most previous NWP convectionallowing precipitation studies, is repeated here.

E	FBIAS by Lead Forecast (h)					
Experiment —	0.25	1.75	3.50	5.00		
HRRR0	1.23	1.36	1.46	1.16		
HRRR3	1.53	1.33	1.18	1.08		
VLAPS0	1.73	1.97	1.44	1.02		
VLAPS3	1.65	1.37	1.27	1.16		
VLAPS3A	2.02	1.98	1.49	1.25		
VLAPS3AO	2.32	2.10	1.42	1.01		
VLAPS3O	2.37	1.55	1.10	0.91		

Table 4. FBIAS for 25 dBZ threshold at four different lead forecast times (h).

Table 5. FBIAS for 35 dBZ threshold at four different lead forecast times (h).

E-m - mires - m t	FBIAS by Lead Forecast (h)				
Experiment —	0.25	1.75	3.50	5.00	
HRRR0	1.08	1.56	1.77	1.48	
HRRR3	1.79	1.53	1.53	1.37	
VLAPS0	1.86	2.40	1.66	1.20	
VLAPS3	1.82	1.55	1.53	1.60	
VLAPS3A	2.41	2.42	1.79	1.69	
VLAPS3AO	2.76	2.49	1.65	1.33	
VLAPS3O	2.67	1.64	1.23	1.25	

Table 6. FSS as it varies across three different neighborhood sizes for the 10 dBZ threshold, and for two different lead forecast times (near the start and end of the forecast period). Note that additional thresholds of 25 dBZ and 35 dBZ are shown only for the 9 km neighborhood size (since 9 km is of interest due to it closely representing the effective resolution of the WRF output produced from the 1 km grid spacing nest).

		FSS by Neighborhood Size (NS), Threshold (T), and Lead Time (LT)						
	NS	11	1 km		9 km		17 km	
	Т	10 c	10 dBZ		10/25/35 dBZ		10 dBZ	
Experiment	LT	0.25 h	5.00 h	0.25 h	5.00 h	0.25 h	5.00 h	
HRRR0		0.43	0.65	0.50/0.22/0.11	0.71/0.42/0.28	0.54	0.74	
HRRR3		0.50	0.61	0.59/0.33/0.20	0.67/0.41/0.28	0.64	0.70	
VLAPS0		0.67	0.56	0.77/0.75/0.72	0.62/0.22/0.11	0.80	0.65	
VLAPS3		0.51	0.61	0.60/0.29/0.18	0.67/0.53/0.34	0.64	0.70	
VLAPS3A		0.54	0.64	0.64/0.44/0.33	0.70/0.41/0.26	0.68	0.72	
VLAPS3AO		0.51	0.55	0.60/0.41/0.31	0.61/0.26/0.11	0.64	0.64	
VLAPS3O		0.45	0.52	0.51/0.28/0.19	0.57/0.46/0.31	0.55	0.59	

4. Discussion

This study's main goal was to complete a preliminary investigation of how nudging and vLAPS 3DVAR compare, and how combining them could improve short-range nowcasting. This paper's focus was more upon the overall methods, and less upon the specific differences across the forecast results, because statistical evaluation of additional cases is needed to draw conclusive statements.

From a subjective standpoint, the experiments captured the main features/structures and general timing of the severe convective outbreak over Oklahoma on the afternoon of 20 May 2013 (some more realistically than others). This was a strongly forced convective

event and was likely to be inherently more predictable [59–62]. The FSS and CSI metrics showed that the experiments performed similarly around the 2 h forecast, while differences among the experiments were somewhat larger at other forecast hours. VLAPS0 is an outlier with much higher scores over the initial lead forecast hour (but generally the lowest scores during the 3 to 5 h lead forecast period). The experiments that applied data assimilation across a short pre-forecast period, including the VLAPS3AO hybrid observation/analysis nudging approach introduced by the authors of this paper, appeared to improve FSS and CSI at the start of the nowcast cycles compared to not performing data assimilation during this pre-forecast period (i.e., VLAPS3AO, VLAPS3A, and VLAPS3O compared to VLAPS3). However, they could not match the FSS and CSI gained by inserting vLAPS analyses at the beginning of the simulation without a pre-forecast period (VLAPS0).

The VLAPS3 combination of starting the model integration from a 3D variational analysis (leveraging radar data), followed by a subsequent 3 h pre-forecast period, performed well relative to most of the other experiments for bias in the first portion of the forecast period (HRRR0 had better bias and HRRR3 was very similar to VLAPS3). It also performed better than all of the other experiments for CSI/FSS in the latter portion of the forecast period.

While analysis nudging toward high-resolution 3DVAR vLAPS analyses shows potential promise for improving short-term convection forecasts, the results suggest further work is needed to best leverage high-resolution analyses. The rapid increase in overforecast bias in VLAPS0 at the beginning of the forecast and CSI and FSS performing worse than any other experiment later in the forecast suggest that the VLAPS0 method of directly inserting the analysis at the beginning of the forecast may not be optimal. Analysis nudging towards a series of vLAPS analyses may be a way to improve the assimilation of vLAPS analyses. However, two factors suggest room for improvement in the methodology used to analysis nudge towards the vLAPS analyses (VLAPS3A, VLAPS3AO). One factor is that the analysis nudging experiments have a noticeably higher overforecast bias than an experiment started with the same initial conditions but no subsequent data assimilation (VLAPS3). The second factor is that during the first hour, directly inserting vLAPS analyses at the beginning of the forecast (VLAPS0) performs much better (in terms of CSI and FSS) than analysis nudging towards the vLAPS analyses. Analysis nudging is not normally used to assimilate analyses with the high temporal (15 min) and horizontal (1 km) spacing used here. Thus, informing the analysis nudging weights by previous research may have resulted in analysis nudging having been applied more weakly or more strongly than would best assimilate the analyses. Additionally, in order to fully assimilate the vLAPS analyses, nudging towards additional fields from the analyses may be needed to supplement those currently available for analysis nudging in WRF. Given the verification results, it is not clear that using observation nudging (VLAPS3A vs. VLAPS3AO) adds values in this case; this may be because the high-temporal and spatial resolution observations used in the vLAPS analysis limit the added value of assimilating individual observations at a time centered on their valid times.

Both HRRR0 and HRRR3 (also in the WS group) proved competitive in FSS and CSI scores after approximately the first 90 min of the lead forecast and out through the longer lead times (and for FBIAS throughout the forecast); this is consistent with both the high-resolution nature and the sophisticated data assimilation methodology built into the HRRR cycling model itself. Since the 15 UTC HRRR cycle forecasts were leveraged in every experiment across all cycle times, the oldest HRRR forecast used to provide a lateral boundary condition (at 00 UTC 21 May) had a 9 h lead time. Some recent studies [63–65] indicate that convection-resolving NWP models, using sophisticated data assimilation approaches (including radar data ingest), can show skill and predictability out to 6–12 h. This is especially true for strongly forced convection, which is the type of convection in this study. Weakly forced convection (such as the type common over the summer in the southeastern US) tends to be considerably less predictable at these same resolutions, due to less certainty in capturing the convective initiation mechanisms [66].

Two advantages our simulations had were that (1) a high-quality 3 km HRRR forecast produced at 15 UTC was available for creating lateral boundary tendencies (and either the initial conditions or the first guess for the analysis used as the initial conditions); and (2) multiple Doppler WSR-88D radars were available to the vLAPS analyses. If we consider a forward-deployed army unit running a similar modeling configuration on a laptop with restrictive communication bandwidth availability, conditions such as (1) and (2) above will likely not typically be present. However, this study shows that atmospheric data assimilation approaches not reliant on a full EnKF [67] or 4DVAR [68] approach, and not in need of high-performance computing cluster solutions, can still provide value-added short-range nowcast guidance to local human forecasters and automated forecast systems and weather decision support tools.

While this study used a single, large domain to simplify the experiment setup (made possible by the availability of the high-resolution HRRR output), in order to make this tractable for a forward-deployed system, nested grids with a much smaller innermost domain would be needed. In the case of a tactical NWP-based nowcasting system (a rapid update cycling model), the ability to collect and effectively assimilate special tactical weather observations (sensors on unmanned aerial systems, lidar, radar, etc.) along with non-traditional sources of weather observations (for example, the Air Force Global Synthetic Weather Radar product [69], which leverages satellite observations and other data sources) will be an important avenue to full success. Additionally, while this study used vLAPS analyses, the techniques investigated could be used with other high-resolution analyses.

In terms of convective forecasting, which was the focus of this paper, there is still ample work remaining towards improvement upon the bias and skill issues, especially at the higher reflectivity threshold levels. Future work should investigate additional cases to explore the generality of the results seen in this case, including exploring whether the high reflectivity bias seen in the current case is seen in other cases and determining the causes of this bias. Given that this was a strongly forced and large-scale convective outbreak, these issues are likely going to be even more challenging in weakly forced convective environments [66]. This case study describes methodologies that can be used to apply 3DVAR analyses in conjunction with nudging, demonstrates that applying these methodologies for vLAPS analyses for this case study does not show a clear overall improvement, and suggests areas for future research to explore how to improve the combination of 3DVAR analyses and nudging. In this case study, directly applying the vLAPS analyses at the 0 h forecast time without analysis nudging resulted in higher FSS and CSI than the other experiments throughout the 0–1 h forecast, but this was also the only experiment with bias increasing substantially during this time period. Similarly, Jiang et al. [19] found that in WRF forecasts initialized with vLAPS, the equitable threat scores were highest at the beginning of the forecasts

At a national level, convection-permitting models are already performing with good skill and with a high degree of sophistication in data assimilation strategies and NWP physics [70]. Running limited-area mesoscale modeling configurations on hardware technology as restricted as a desktop or laptop, while maintaining a flexibility to adjust quickly for operations over diverse and often data-restricted locations around the world, will require making intelligent adjustments to the national approaches at convection-permitting scales. The 1 km (and even finer) grid spacing is important for resolving many boundary layer flow phenomena that impact both military and civilian interests near the earth's surface, and convective forecasting is just one aspect of those (although a potentially highly impactful one).

Author Contributions: Conceptualization, B.P.R., H.C. and R.E.D.; methodology, B.P.R., H.C., R.E.D., Y.X. and S.A.; software, B.P.R., Y.X. and S.A.; validation, H.C. and J.W.R.; formal analysis, H.C. and J.W.R.; investigation, B.P.R., R.E.D., Y.X. and S.A.; writing—original draft preparation, B.P.R. and R.E.D.; writing—review and editing, B.P.R., H.C., R.E.D., Y.X., S.A. and J.W.R.; visualization, B.P.R. and H.C. All authors have read and agreed to the published version of the manuscript.

Funding: The US Army DEVCOM Army Research Laboratory (ARL) authors (Brian Reen, Huaqing Cai, Robert Dumais, and John Raby) were funded by ARL. Yuanfu Xie and Steve Albers were partially funded by ARL and partially funded by NOAA/GSD (and Steve Albers by CIRA).

Data Availability Statement: MADIS observations for assimilation were downloaded from https: //madis-data.ncep.noaa.gov/ (accessed on 10 November 2021). Pilot observations, WSR-88D radar data, and GOES data were obtained from NOAA Global Systems Laboratory (GSL). The WRF verification against radar reflectivity that is the basis for the Results section is available at https: //doi.org/10.6084/m9.figshare.14818497.v1 (accessed on 10 November 2021). The remaining code and datasets have not been approved by the authors' institutions for public release.

Acknowledgments: Hongli Jiang (formerly NOAA ESRL and CIRA) contributed in various ways to this study including providing guidance on using vLAPS analyses in WRF and running an earlier version of some of the experiments. The Real-Time Mesoscale Analysis use/reject lists provided by Steve Levine at the National Centers for Environmental Prediction's Environmental Modeling Center greatly facilitated making full use of the MADIS observational dataset. Cindy Bruyère (NCAR) provided the modified version of Ungrib used in this work to improve quality control of the observations for nudging. John Halley Gotway (NCAR) provided a script that facilitated ingestion of vLAPS-processed radar data into the Model Evaluation Tools (MET) for verification. Jeffrey Smith (ARL) assisted with file processing related to verification. This work was supported, in part, by high-performance computer time and resources from the DoD High Performance Computing Modernization Program. This study was made possible, in part, due to the data made available to NOAA by various data providers for inclusion in MADIS. MET was used in this study and was developed at NCAR through grants from the National Science Foundation, NOAA, the United States (US) Air Force, and the US Department of Energy. NCAR is sponsored by the US National Science Foundation.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Hu, M.; Xue, M.; Brewster, K. 3DVAR and Cloud Analysis with WSR-88D Level-II Data for the Prediction of the Fort Worth, Texas, Tornadic Thunderstorms. Part I: Cloud Analysis and Its Impact. *Mon. Weather Rev.* **2006**, *134*, 675–698. [CrossRef]
- Xiao, Q.; Sun, J. Multiple-Radar Data Assimilation and Short-Range Quantitative Precipitation Forecasting of a Squall Line Observed during IHOP_2002. *Mon. Weather Rev.* 2007, 135, 3381–3404. [CrossRef]
- 3. Schroeder, A.J.; Stauffer, D.R.; Seaman, N.L.; Deng, A.; Gibbs, A.M.; Hunter, G.K.; Young, G.S. An Automated High-Resolution, Rapidly Relocatable Meteorological Nowcasting and Prediction System. *Mon. Weather Rev.* **2006**, *134*, 1237–1265. [CrossRef]
- Reen, B.P.; Stauffer, D.R. Data assimilation strategies in the planetary boundary layer. *Bound.-Layer Meteorol.* 2010, 137, 237–269. [CrossRef]
- Hu, M.; Benjamin, S.G.; Ladwig, T.T.; Dowell, D.C.; Weygandt, S.S.; Alexander, C.R.; Whitaker, J.S. GSI Three-Dimensional Ensemble–Variational Hybrid Data Assimilation Using a Global Ensemble for the Regional Rapid Refresh Model. *Mon. Weather Rev.* 2017, 145, 4205–4225. [CrossRef]
- Lee, M.; Kuo, Y.; Barker, D.M.; Lim, E. Incremental Analysis Updates Initialization Technique Applied to 10-km MM5 and MM5 3DVAR. *Mon. Weather Rev.* 2006, 134, 1389–1404. [CrossRef]
- Brewster, K.A.; Stratman, D.R. Tuning an analysis and incremental analysis updating assimilation for an efficient high-resolution forecast system. In Proceedings of the 20th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans and Land Surface (IAOS-AOLS), New Orleans, LA, USA, 11–14 January 2016; Available online: https://ams.confex.com/ ams/96Annual/webprogram/Paper289235.html (accessed on 10 November 2021).
- 8. Lorenc, A.C.; Rawlins, F. Why does 4D-Var beat 3D-Var? Quart. J. Roy. Meteor. Soc. 2005, 131, 3247–3257. [CrossRef]
- 9. Stauffer, D.R.; Seaman, N.L. Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I: Experiments with Synoptic-Scale Data. *Mon. Weather Rev.* **1990**, *118*, 1250–1277. [CrossRef]
- Stauffer, D.R.; Seaman, N.L.; Binkowski, F.S. Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model Part II: Effects of Data Assimilation within the Planetary Boundary Layer. *Mon. Weather Rev.* 1991, 119, 734–754. [CrossRef]
- Shaw, B.L.; Carpenter, R.L.; Spencer, P.L.; DuFran, Z. Commercial Implementation of WRF with Efficient Computing and Advanced Data Assimilation. In Proceedings of the 9th Annual WRF User's Workshop, Boulder, CO, USA, 23–27 June 2008; Available online: http://www2.mmm.ucar.edu/wrf/users/workshops/WS2008/abstracts/2-05.pdf (accessed on 10 November 2021).
- 12. Liu, Y.; Yu, W.; Warner, T.; Sun, J.; Xiao, Q.; Pace, J. Short-term explicit forecasting and nowcasting of convection systems with WRF using a hybrid radar data assimilation approach. In Proceedings of the 23rd Conference on Weather Analysis and Forecasting/19th Conference on Numerical Weather Prediction, Omaha, NE, USA, 1–5 June 2009; Available online: https://ams.confex.com/ams/23WAF19NWP/webprogram/Paper154264.html (accessed on 10 November 2021).

- 13. Lei, L.; Stauffer, D.R.; Haupt, S.E.; Young, G.S. A hybrid nudging- ensemble Kalman filter approach to data assimilation. Part I: Application in the Lorenz system. *Tellus* **2012**, *64A*, 18484. [CrossRef]
- 14. Lei, L.; Stauffer, D.R.; Deng, A. A hybrid nudging-ensemble Kalman filter approach to data assimilation. Part II: Application in a shallow-water model. *Tellus* **2012**, *64A*, 18485. [CrossRef]
- Lei, L.; Stauffer, D.R.; Deng, A. A hybrid nudging-ensemble Kalman filter approach to data assimilation in WRF/DART. Q. J. R. Meteorol. Soc. 2012, 138, 2066–2078. [CrossRef]
- Skamarock, W.C.; Coauthors. A Description of the Advanced Research WRF Version 3; NCAR Technical Note NCAR/TN-4751STR; National Center for Atmospheric Research: Boulder, CO, USA, 2008; p. 113. Available online: http://dx.doi.org/10.5065/D68S4 MVH (accessed on 10 November 2021).
- Liu, Y.; Wu, Y.; Pan, L.; Bourgeois, A.; Roux, G.; Knievel, J.C.; Hacker, J.; Pace, J.; Gallagher, W.; Halvorson, S.F. Recent developments of NCAR's 4D-Relaxation Ensemble Kalman Filter system. In Proceedings of the 19th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS), Phoenix, AZ, USA, 5–8 January 2015; Available online: https://ams.confex.com/ams/95Annual/webprogram/Paper269311.html (accessed on 10 November 2021).
- Lei, L.; Hacker, J.P. Nudging, Ensemble, and Nudging Ensembles for Data Assimilation in the Presence of Model Error. *Mon. Weather Rev.* 2015, 143, 2600–2610. [CrossRef]
- Jiang, H.; Albers, S.; Xie, Y.; Toth, Z.; Jankov, I.; Scotten, M.; Picca, J.; Stumpf, G.; Kingfield, D.; Birkenheuer, D.; et al. Real-Time Applications of the Variational Version of the Local Analysis and Prediction System (vLAPS). *Bull. Am. Meteorol. Soc.* 2015, *96*, 2045–2057. [CrossRef]
- Reen, B.P.; Xie, Y.; Cai, H.; Albers, S.; Dumais, R.E., Jr.; Jiang, H. Incorporating Variational Local Analysis and Prediction System (vLAPS) Analyses with Nudging Data Assimilation: Methodology and Initial Results; ARL-TR-8145; U.S. Army Research Laboratory: Adelphi, MD, USA, 2017; p. 76. Available online: https://apps.dtic.mil/sti/pdfs/AD1039383.pdf (accessed on 10 November 2021).
- Rogers, R.E.; Deng, A.; Stauffer, D.R.; Gaudet, B.J.; Jia, Y.; Soong, S.; Tanrikulu, S. Application of the Weather Research and Forecasting Model for Air Quality Modeling in the San Francisco Bay Area. J. Appl. Meteorol. Climatol. 2013, 52, 1953–1973. [CrossRef]
- 22. Expósito, F.J.; González, A.; Pérez, J.C.; Díaz, J.P.; Taima, D. High-Resolution Future Projections of Temperature and Precipitation in the Canary Islands. J. Clim. 2015, 28, 7846–7856. [CrossRef]
- 23. Huo, Z.; Liu, Y.; Wei, M.; Shi, Y.; Fang, C.; Shu, Z.; Li, Y. Hydrometeor and Latent Heating Nudging for Radar Reflectivity Assimilation: Response to the Model States and Uncertainties. *Remote Sens.* **2021**, *13*, 3821. [CrossRef]
- Janjić, Z.I. Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso Model; NCEP Office Note 437; National Centers for Environmental Prediction: Camp Springs, MD, USA, 2001; p. 61. Available online: http://www.lib.ncep. noaa.gov/ncepofficenotes/files/on437.pdf (accessed on 10 November 2021).
- Lee, J.A.; Kolczyinski, W.C.; McCandless, T.C.; Haupt, S.E. An Objective Methodology for Configuring and Down-Selecting an NWP Ensemble for Low-Level Wind Prediction. *Mon. Weather Rev.* 2012, 140, 2270–2286. [CrossRef]
- Reen, B.P.; Schmehl, K.J.; Young, G.S.; Lee, J.A.; Haupt, S.E.; Stauffer, D.R. Uncertainty in Contaminant Concentration Fields Resulting from Atmospheric Boundary Layer Depth Uncertainty. J. Appl. Meteorol. Climatol. 2014, 53, 2610–2626. [CrossRef]
- Thompson, G.; Field, P.R.; Rasmussen, R.M.; Hall, W.D. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.* 2008, 136, 5095–5115. [CrossRef]
- 28. Mlawer, E.J.; Taubman, S.J.; Brown, P.D.; Iacono, M.J.; Clough, S.A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.* **1997**, *102*, 16663–16682. [CrossRef]
- 29. Dudhia, J. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. J. Atmos. Sci. 1989, 46, 3077–3107. [CrossRef]
- 30. Chen, F.; Dudhia, J. Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Weather Rev.* 2001, 129, 569–585. [CrossRef]
- McGinley, J.A.; Albers, S.C.; Stamus, P.A. Validation of a Composite Convective Index as Defined by a Real-Time Local Analysis System. Weather Forecast. 1991, 6, 337–356. [CrossRef]
- 32. Hiemstra, C.A.; Liston, G.E.; Pielke, R.A.; Birkenheuer, D.L.; Albers, S.C. Comparing Local Analysis and Prediction System (LAPS) Assimilations with Independent Observations. *Weather Forecast.* **2006**, *21*, 1024–1040. [CrossRef]
- 33. Albers, S.C.; McGinley, J.A.; Birkenheuer, D.L.; Smart, J.R. The Local Analysis and Prediction System (LAPS): Analyses of Clouds, Precipitation, and Temperature. *Weather Forecast.* **1996**, *11*, 273–287. [CrossRef]
- 34. Xie, Y.; Koch, S.; McGinley, J.; Albers, S.; Bieringer, P.E.; Wolfson, M.; Chan, M. A Space–Time Multiscale Analysis System: A Sequential Variational Analysis Approach. *Mon. Weather Rev.* **2011**, *139*, 1224–1240. [CrossRef]
- Xie, Y.F.; MacDonald, A.E. Selection of momentum variables for a three-dimensional variational analysis. *Pure Appl. Geophys.* 2012, 169, 335–351. [CrossRef]
- 36. Xie, Y.; Lu, C.; Browning, G.L. Impact of Formulation of Cost Function and Constraints on Three-Dimensional Variational Data Assimilation. *Mon. Weather Rev.* 2002, 130, 2433–2447. [CrossRef]

- 37. Alexander, C.; Smirnova, T.; Hu, M.; Weygandt, S.; Dowell, D.; Olson, J.; Benjamin, S.; James, E.; Brown, J.; Hofmann, P.; et al. Recent advancements in convective-scale storm prediction with the coupled Rapid Refresh (RAP)/High-Resolution Rapid Refresh (HRRR) forecast system. In Proceedings of the 14th Annual WRF Users' Workshop, Boulder, CO, USA, 24–28 June 2013; Available online: http://www2.mmm.ucar.edu/wrf/users/workshops/WS2013/ppts/2.3.pdf (accessed on 10 November 2021).
- Stauffer, D.R.; Seaman, N.L. Multiscale Four-Dimensional Data Assimilation. J. Appl. Meteorol. Climatol. 1994, 33, 416–434. [CrossRef]
- Gilliam, R.C.; Godowitch, J.M.; Rao, S.T. Improving the horizontal transport in the lower troposphere with four dimensional data assimilation. *Atmos. Environ.* 2012, 53, 186–201. [CrossRef]
- Reen, B.P. A Brief Guide to Observation Nudging in WRF. 2016. Available online: http://www2.mmm.ucar.edu/wrf/users/ docs/ObsNudgingGuide.pdf (accessed on 10 November 2021).
- De Pondeca, M.S.; Manikin, G.S.; DiMego, G.; Benjamin, S.G.; Parrish, D.F.; Purser, R.J.; Wu, W.; Horel, J.D.; Myrick, D.T.; Lin, Y.; et al. The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current Status and Development. *Weather Forecast.* 2011, 26, 593–612. [CrossRef]
- National Center for Atmospheric Research. User's Guide for the Advanced Research WRF (ARW) Modeling System Version 3.8. 2016. Available online: https://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/user_guide_V3.8/contents.html (accessed on 10 November 2021).
- 43. Reen, B.P.; Dumais, R.E.; Passner, J.E. Mitigating Excessive Drying from the Use of Observations in Mesoscale Modeling. *J. Appl. Meteorol. Climatol.* **2016**, *55*, 365–388. [CrossRef]
- Atkins, N.T.; Butler, K.M.; Flynn, K.R.; Wakimoto, R.M. An Integrated Damage, Visual, and Radar Analysis of the 2013 Moore, Oklahoma, EF5 Tornado. Bull. Am. Meteorol. Soc. 2014, 95, 1549–1561. [CrossRef]
- 45. Burgess, D.; Ortega, K.; Stumpf, G.; Garfield, G.; Karstens, C.; Meyer, T.; Smith, B.; Speheger, D.; Ladue, J.; Smith, R.; et al. 20 May 2013 Moore, Oklahoma, Tornado: Damage Survey and Analysis. *Weather Forecast.* **2014**, *29*, 1229–1237. [CrossRef]
- Kurdzo, J.M.; Bodine, D.J.; Cheong, B.L.; Palmer, R.D. High-Temporal Resolution Polarimetric X-Band Doppler Radar Observations of the 20 May 2013 Moore, Oklahoma, Tornado. *Mon. Weather Rev.* 2015, 143, 2711–2735. [CrossRef]
- Hanley, K.E.; Barrett, A.I.; Lean, H.W. Simulating the 20 May 2013 Moore, Oklahoma tornado with a 100-metre grid-length NWP model. *Atmos. Sci. Lett.* 2016, 17, 453–461. [CrossRef]
- 48. Zhang, Y.; Zhang, F.; Stensrud, D.J.; Meng, Z. Practical Predictability of the 20 May 2013 Tornadic Thunderstorm Event in Oklahoma: Sensitivity to Synoptic Timing and Topographical Influence. *Mon. Weather Rev.* 2015, 143, 2973–2997. [CrossRef]
- 49. Zhang, Y.; Zhang, F.; Stensrud, D.J.; Meng, Z. Intrinsic Predictability of the 20 May 2013 Tornadic Thunderstorm Event in Oklahoma at Storm Scales. *Mon. Weather Rev.* 2016, 144, 1273–1298. [CrossRef]
- Snook, N.; Jung, Y.; Brotzge, J.; Putnam, B.; Xue, M. Prediction and Ensemble Forecast Verification of Hail in the Supercell Storms of 20 May 2013. Weather Forecast. 2016, 31, 811–825. [CrossRef]
- Skok, G.; Roberts, N. Analysis of fractions skill score properties for random precipitation fields and ECMWF forecasts. Q. J. R. Meteorol. Soc. 2016, 142, 2599–2610. [CrossRef]
- 52. Schaefer, J.T. The critical success index as an indicator of warning skill. Weather Forecast. 1990, 5, 570–575. [CrossRef]
- 53. Wilks, D.S. Statistical Methods in the Atmospheric Sciences, 2nd ed.; Academic Press: San Diego, CA, USA, 2005; p. 264.
- 54. Hong, S.-Y.; Park, H.; Cheong, H.-B.; Kim, J.-E.E.; Koo, M.-S.; Jang, J.; Ham, S.; Hwang, S.-O.; Park, B.-K.; Chang, E.-C.; et al. The global/regional integrated model system (GRIMs). *Asia Pac. J. Atmos. Sci.* **2013**, *49*, 219–243. [CrossRef]
- 55. Sun, J.; Xue, M.; Wilson, J.W.; Zawadzki, I.; Ballard, S.P.; Onvlee-Hooimeyer, J.; Joe, P.; Barker, D.M.; Li, P.-W.; Golding, B.; et al. Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Am. Meteorol. Soc.* 2014, 95, 409–426. [CrossRef]
- 56. Koch, S.; Ferrie, B.; Stoelinga, M.; Szoke, E.; Weiss, J.; Kain, S.J. The use of simulated radar reflectivity fields in the diagnosis of mesoscale phenomena from high-resolution WRF model forecasts. In Proceedings of the 11th Conference on Mesoscale Processes/32nd Conference on Radar Meteorology, Albuquerque, NM, USA, 24–28 October 2005; Available online: https://ams.confex.com/ams/32Rad11Meso/techprogram/paper_97032.htm (accessed on 10 November 2021).
- Skamarock, W.C.; Park, S.-H.; Klemp, J.B.; Snyder, C. Atmospheric kinetic energy spectra from global high-resolution nonhydrostatic simulations. *J. Atmos. Sci.* 2014, 71, 4369–4381. [CrossRef]
- 58. Shaw, B.L.; Albers, S.; Birkenheuer, D.; Brown, J.; McGinley, J.A.; Schultz, P.; Smart, J.R.; Szoke, E. Application of the Local Analysis and Prediction System (LAPS) diabatic initialization of mesoscale numerical weather prediction models for the IHOP—2002 field experiment. In Proceedings of the 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction, Seattle, WA, USA, 12–15 January 2004; Available online: https://ams.confex.com/ams/84Annual/techprogram/ paper_69700.htm (accessed on 10 November 2021).
- 59. Wapler, K.; Harnisch, F.; Pardowitz, T.; Senf, F. Characterisation and predictability of a strong and a weak forcing severe convective event a multi-data approach. *Meteorol. Z.* **2015**, *24*, 393–410. [CrossRef]
- 60. Keil, C.; Heinlein, F.; Craig, G.C. The convective adjustment time-scale as indicator of predictability of convective precipitation. *Q. J. R. Meteorol. Soc.* **2014**, *140*, 480–490. [CrossRef]
- 61. Stensrud, D.J.; Fritsch, J.M. Mesoscale convective systems in weakly forced large-scale environments. Part II: Generation of a mesoscale initial condition. *Mon. Weather Rev.* **1994**, 122, 2068–2083. [CrossRef]

- 62. Duda, J.D.; Gallus, W.A. The impact of large-scale forcing on skill of simulated convective initiation and upscale evolution with convection-allowing grid spacings in the WRF. *Weather Forecast.* **2013**, *28*, 994–1018. [CrossRef]
- 63. Stratman, D.R.; Coniglio, M.C.; Koch, S.E.; Xue, M. Use of Multiple Verification Methods to Evaluate Forecasts of Convection from Hot- and Cold-Start Convection-Allowing Models. *Weather Forecast.* **2013**, *28*, 119–138. [CrossRef]
- 64. Carlberg, B.R.; Gallus, W.A.; Franz, K.J. A Preliminary Examination of WRF Ensemble Prediction of Convective Mode Evolution. *Weather Forecast.* **2018**, *33*, 783–798. [CrossRef]
- 65. Pan, L.; Liu, Y.; Lei, Y.; Xu, M.; Child, P.; Jacobs, N. Development of a CONUS radar data assimilation WRF-RTFDDA system for convection-resolvable analysis and prediction. In Proceedings of the 16th WRF Annual User's Workshop, Boulder, CO, USA, 15–19 June 2015; Available online: https://www2.mmm.ucar.edu/wrf/users/workshops/WS2015/WorkshopPapers.php (accessed on 10 November 2021).
- Zeng, Y.; Janjić, T.; de Lozar, A.; Blahak, U.; Reich, H.; Keil, C.; Seifert, A. Representation of model error in convective-scale data assimilation: Additive noise, relaxation methods, and combinations. J. Adv. Model. Earth Syst. 2018, 10, 2889–2911. [CrossRef]
- 67. James, E.P.; Benjamin, S.G. Observation System Experiments with the Hourly Updating Rapid Refresh Model Using GSI Hybrid Ensemble-Variational Data Assimilation. *Mon. Weather Rev.* **2017**, *145*, 2897–2918. [CrossRef]
- 68. Ban, J.; Liu, Z.; Zhang, X.; Huang, X.-Y.; Wang, H. Precipitation data assimilation in WRFDA 4D-Var: Implementation and application to convection-permitting forecasts over United States. *Tellus* **2017**, *69*, 1368310. [CrossRef]
- 69. Veillette, M.S.; Iskenderian, H.; Lamey, P.M.; Mattioli, C.J.; Banerjee, A.; Worris, M.; Porschitsky, A.B.; Ferris, R.F.; Manwelyan, A.; Rajogoplalan, S.; et al. Global Synthetic Weather Radar in AWS GovCloud for the U.S. Air Force. In Proceedings of the 19th Conference on Artificial Intelligence for Environmental Science, Boston, MA, USA, 13–16 January 2020; Available online: https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/363150 (accessed on 10 November 2021).
- 70. Bytheway, J.L.; Kummerow, C.D. Toward an object-based assessment of high-resolution forecasts of long-lived convective precipitation in the central U.S. *J. Adv. Model. Earth Syst.* **2015**, *7*, 1248–1264. [CrossRef]