



ARL-TR-7881 • Nov 2016



Domain-Level Assessment of the Weather Running Estimate–Nowcast (WRE–N) Model

by John W Raby and Jeffrey A Smith

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.



Domain-Level Assessment of the Weather Running Estimate–Nowcast (WRE–N) Model

by John W Raby and Jeffrey A Smith

Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) November 2016		2. REPORT TYPE Technical Report		3. DATES COVERED (From – To) October 2015 to September 2016	
4. TITLE AND SUBTITLE Domain-Level Assessment of the Weather Running Estimate–Nowcast (WRE–N) Model				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) John W Raby and Jeffrey A Smith				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory Computational and Information Sciences Directorate ATTN: RDRL-CIE-M White Sands Missile Range, NM 88002				8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-7881	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The Weather Research and Forecasting Model (WRF) is a numerical weather-prediction model that has been used for many applications including My Weather Impacts Decision Aid. WRF is maintained by the National Center for Atmospheric Research, which has developed a suite of Model Evaluation Tools (MET) to evaluate the accuracy of WRF forecasts. In this technical report, we discuss our use of the MET to assess the domain-level errors of the Weather Running Estimate–Nowcast (WRE–N) model. In addition, the domain-level errors were calculated for the Global Forecast System (GFS) model, which is used to initialize the WRE–N. The output used for this study was generated using 3 different configurations of the WRE–N, with and without observation assimilation, that were run over 2 different triple-nested domains centered near San Diego, California. We selected 5 case-study days in February–March 2012 with varied weather conditions. The results of the study suggest the observation assimilation improves the forecast under certain conditions. The results show the WRE–N generally performs better than the GFS model with some limitations. More comprehensive verification studies are needed to conclusively determine the value added by varying configurations of the WRE–N.					
15. SUBJECT TERMS model verification, Model Evaluation Tools, error statistics, assessment, Numerical Weather Prediction, nowcast, Weather Running Estimate, validation					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 46	19a. NAME OF RESPONSIBLE PERSON John W Raby
REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) 575-678-2004

Contents

List of Figures	iv
List of Tables	v
Preface	vi
Acknowledgments	vii
Executive Summary	ix
1. Introduction and Background	1
2. Domain and Model	2
2.1 Observations	4
2.2 Parameterizations	4
2.3 Case-Study Days	5
3. Data Preparation Using MET	6
4. Data Analysis	7
4.1 Value Added to WRE–N by FDDA at 1.75-km Grid Spacing	7
4.2 Value Added to WRE–N by FDDA at 5.25- and 1.75-km Grid Spacing	10
4.3 Value Added by Decreased Grid Spacing	14
4.4 Performance Comparison of 2 WRE–N Configurations	18
4.5 Performance Comparison: Dumais WRE–N with FDDA vs. the GFS	23
5. Conclusions and Final Comments	26
6. References	28
List of Symbols, Abbreviations, and Acronyms	32
Distribution List	34

List of Figures

Fig. 1	Two triple-nested model domains; each domain's center is coincident and both nested configurations are centered near San Diego, California (Google Earth 2016).....	3
Fig. 2	Bias and RMSE errors for WRE-N with and without FDDA for 2-m-AGL TMP (in degrees Kelvin [K])	7
Fig. 3	Bias and RMSE errors for WRE-N with and without FDDA for 2-m-AGL DPT (K)	8
Fig. 4	Bias and RMSE errors for WRE-N with and without FDDA for 10-m-AGL U (m/s).....	9
Fig. 5	Bias and RMSE errors for WRE-N with and without FDDA for 10-m-AGL V (m/s).....	10
Fig. 6	Bias and RMSE errors for WRE-N with and without FDDA for 2-m-AGL TMP (K)	11
Fig. 7	Bias and RMSE errors for WRE-N with and without FDDA for 2-m-AGL DPT (K).....	12
Fig. 8	Bias and RMSE errors for WRE-N with and without FDDA for 10-m-AGL U (m/s).....	13
Fig. 9	Bias and RMSE errors for WRE-N with and without FDDA for 10-m-AGL V (m/s).....	14
Fig. 10	Bias and RMSE errors for the 3 grids for Dumais and Passner WRE-N with FDDA for 2-m-AGL TMP (K)	15
Fig. 11	Bias and RMSE errors for the 3 grids for Dumais and Passner WRE-N with FDDA for 2-m-AGL DPT (K)	16
Fig. 12	Bias and RMSE errors for the 3 grids for Dumais and Passner WRE-N with FDDA for 10-m-AGL U (m/s).....	17
Fig. 13	Bias and RMSE errors for the 3 grids for Dumais and Passner WRE-N with FDDA for 10-m-AGL V (m/s).....	18
Fig. 14	Bias and RMSE errors for the Reen and Passner 1-km WRE-N with FDDA for 2-m-AGL TMP (K)	19
Fig. 15	Bias and RMSE errors for the Reen and Passner 1-km WRE-N with FDDA for 2-m-AGL DPT (K).....	20
Fig. 16	Bias and RMSE errors for the Reen and Passner 1-km WRE-N with FDDA for 10-m-AGL U (m/s)	21
Fig. 17	Bias and RMSE errors for the Reen and Passner 1-km WRE-N with FDDA for 10-m-AGL V (m/s)	22
Fig. 18	Bias and RMSE errors for the Dumais WRE-N with FDDA and the GFS for 2-m-AGL TMP (K).....	23

Fig. 19	Bias and RMSE errors for the Dumais WRE–N with FDDA and the GFS for 2-m-AGL RH (in %).....	24
Fig. 20	Bias and RMSE errors for the Dumais WRE–N with FDDA and the GFS for 10-m-AGL U (m/s)	25
Fig. 21	Bias and RMSE errors for the Dumais WRE–N with FDDA and the GFS for 10-m-AGL V (m/s)	26

List of Tables

Table 1	WRE-N triple-nested domain dimensions	3
Table 2	WRE-N configurations.....	5
Table 3	Synoptic conditions for the case-study days considered.....	6
Table 4	WRE–N configuration differences.....	19

Preface

This technical report relates to previous works that explored the use of spatial verification techniques with a grid-to-grid approach for the assessment of the Weather Running Estimate–Nowcast (WRE–N) model. However, this report describes the results obtained when a traditional, grid-to-point approach was used to verify the WRE–N model. Portions of this report’s content appeared in ARL-TR-7849¹ and, before it, in ARL-TR-7751.²

¹ Raby JW. Application of a fuzzy verification technique for assessment of the Weather Running Estimate–Nowcast (WRE–N) model. White Sands Missile Range (NM): Army Research Laboratory (US); 2016 Oct. Report No.: ARL-TR-7849.

² Raby JW, Cai H. Verification of spatial forecasts of continuous meteorological variables using categorical and object-based methods. White Sands Missile Range (NM): Army Research Laboratory (US); 2016 Aug. Report No.: ARL-TR-7751.

Acknowledgments

We offer our thanks to Mr Robert Dumais, Mr Jeff Passner, and Dr Brian Reen of the US Army Research Laboratory (ARL). Each of these individuals contributed data and information without which the study could not have been completed.

Many thanks go to Mr Martin Kufus of the ARL Technical Publishing office at White Sands Missile Range, New Mexico, for his consistently high standard of editing.

INTENTIONALLY LEFT BLANK.

Executive Summary

The Weather Research and Forecasting Model (WRF) is a numerical weather-prediction model that has been used for many applications, including My Weather Impacts Decision Aid. WRF is maintained by the National Center for Atmospheric Research, which has developed a suite of Model Evaluation Tools (MET) to evaluate the accuracy of WRF forecasts using observations of meteorological variables such as temperature, relative humidity, and wind.

In this technical report, we discuss our use of the MET to assess the errors aggregated over selected domains of the Weather Running Estimate–Nowcast (WRE–N) model, which uses the Advanced Research WRF (commonly referred to as WRF–ARW) as its core. In addition, the domain-level errors were calculated for the Global Forecast System (GFS) model that is used to initialize the WRE–N. The WRE–N was developed to address the US Army requirement to provide high-resolution weather forecasting to resolve atmospheric features with wavelengths on the order of 5 km or less. This requires models that operate on a model grid spacing of about 1 km or less in the finest, or most resolved, domain.

The output used for this study was generated using 3 different configurations of the WRE–N that were run over 2 different triple-nested domains centered near San Diego, California. The grid spacing of the WRE–N was largest in the outer domain and smallest in the innermost domain. The San Diego area contains a mixture of urban, suburban, agricultural, and mountainous terrain types along with a rich array of observational data with which to conduct domain-level assessments. We selected 5 case-study days in February–March 2012 with varied synoptic weather conditions.

In this study we generated the traditional grid-to-point, continuous error statistics for each of the domains separately to maximize the number of observations used in the domain-level statistics for the 3 model grids. We also scored the 3 different grids of the WRE–N over the innermost domain to compare the errors for each grid spacing.

In this study, we also compared the error statistics for the WRE–N run with the Four-Dimensional Data Assimilation (FDDA) technique that incorporates observations into the model with runs that did not use the FDDA to evaluate the value added by the technique in terms of the error statistics. The error statistics for 2 different configurations of the WRE–N run over the same triple nest were compared.

The results of the study suggest the observation assimilation improves the forecast under certain conditions, but the value added of smaller grid spacing cannot be determined. The results show the WRE-N generally performs better than the GFS model with some limitations. More comprehensive verification studies are needed to conclusively determine the value added by varying configurations of the WRE-N.

1. Introduction and Background

As computing technology has advanced, the weather forecasting task—once the primary role of a human forecaster in theater—has shifted to computerized Numerical Weather Prediction (NWP) models. Scientists around the world have used the Weather Research and Forecasting model (WRF) extensively for many applications (Raby 2016; Raby and Cai 2016). In this study, we have used the Advanced Research version of WRF (Skamarock et al. 2008) that we abbreviate as WRF-ARW. The WRF-ARW includes Four-Dimensional Data Assimilation (FDDA) techniques that can be used to incorporate observations into the model so that forecast quality is improved (Stauffer and Seaman 1994; Deng et al. 2009). The US Army Research Laboratory (ARL) uses WRF-ARW as the core of its Weather Running Estimate-Nowcast (WRE-N) weather-forecasting model.

The Army requires high-resolution weather forecasting to model atmospheric features with wavelengths on the order of 5 km or less, which imposes a requirement for NWP to operate on a model grid spacing on the order of 1 km or less in the finest, or most resolved, domain in order to resolve weather phenomena of interest to the Soldier in theater. The atmospheric flows of interest to the Army include mountain/valley breezes, sea breezes, and other flows induced by differences in land-surface characteristics. High-resolution NWP forecasts need to be validated against observations before their outputs can be used by applications such as the My Weather Impacts Decision Aid (MyWIDA) developed by Brandt et al. (2013). Weather-forecast validation has always been of interest to the civilian and military weather-forecasting community; see, for example, the reviews by Ebert et al. (2013) and Casati et al. (2008) or the guides by Jolliffe and Stephenson (2011) or Wilks (2011). The validation of the models, especially high-resolution NWP, has proven to be especially difficult when addressing small temporal and spatial scales (NRC 2010) that characterize NWP for use in Army applications.

The WRF model is maintained by the National Center for Atmospheric Research (NCAR), which has also developed a suite of Model Evaluation Tools (MET) (NCAR 2013) to evaluate WRF-ARW performance. MET was developed at NCAR through a grant from the US Air Force's 557th Weather Wing (formerly the Air Force Weather Agency). NCAR is sponsored by the National Science Foundation. MET Point-Stat performs traditional grid-to-point verification, while MET Grid-Stat performs grid-to-grid neighborhood verification to account for the uncertainty inherent in high-resolution forecasting. MET Method for Object-based Diagnostic Evaluation has been used to develop techniques for object-based spatial verification of high-resolution forecast grids of meteorological variables.

ARL has employed MET Point-Stat in prior assessments such as that of Raby et al. (2012), who evaluated 2 models to arrive at domain-level conclusions about the various strengths and weaknesses of these models and their accuracies. Point-Stat proved useful as an assessment tool for the 2 models over a regional domain; plans were formulated to expand its use to perform comparisons of various configurations and grid spacings of the WRE-N. The WRE-N was run by 3 modelers (A-Reen, B-Passner, and C-Dumais) using 3 different sets of parameterizations over 2 different sets of triple-nested domains, each at a different grid spacing, with and without FDDA, which provided a robust data set of model output for analysis. In addition, Point-Stat was run on the half-degree-spacing Global Forecast System (GFS) initialization grids used for the WRE-N to provide data for error intercomparisons.

This study provides domain-level assessments of the WRE-N and the GFS initialization grid; these give insight into the value added by the WRE-N, the FDDA, and reduced grid spacing in terms of traditional error statistics as well as characterizations of the temporal variability of these statistics throughout the model's time domain by intercomparing differing model configurations to reveal relative strengths and weaknesses. Our study domain was located in the San Diego, California, region, which we chose because of its varied terrain and large number of weather-observing stations. We chose 5 case-study days with varied synoptic conditions in February-March 2012.

2. Domain and Model

The ARL WRE-N (Dumais et al. 2004; Dumais et al. 2013) has been designed as a convection-allowing application of the WRF-ARW model (Skamarock et al. 2008) with an observation-nudging FDDA option (Deng et al. 2009; Liu et al. 2005). For this investigation, the WRE-N was configured to run over 2 multinest sets of domains to produce a fine inner mesh with 1- and 1.75-km grid spacing; also, it leveraged an external global model for cold-start initial conditions and time-dependent lateral boundary conditions for the outermost nest. Table 1 describes the dimensions of the 2 sets of domains. This global model for ARL development and testing has been the National Center for Environmental Prediction's GFS model (EMC 2003). The WRE-N is envisioned to be a rapid-update, cycling application of WRF-ARW with FDDA and optimally could refresh itself at intervals up to hourly, dependent upon the observation network (Dumais and Reen 2013).

Table 1 WRE-N triple-nested domain dimensions

Modeler	East–West dimension (km)	North–South dimension (km)	Grid spacing (km)
Reen, Passner	1575	1575	9
	723	723	3
	127	127	1
Dumais	1780	1780	15.75
	761	761	5.25
	506	506	1.75

For this study, the model runs had a base time of 1200 coordinated universal time (UTC) and produced output for each hour from 1200 UTC to 0600 UTC of the following day for a total of 19 model outputs used on each of 5 days in February–March 2012. Figure 1 depicts the modeling domains. The domains outlined in yellow are for the Dumais model and in red for the Reen and Passner models.

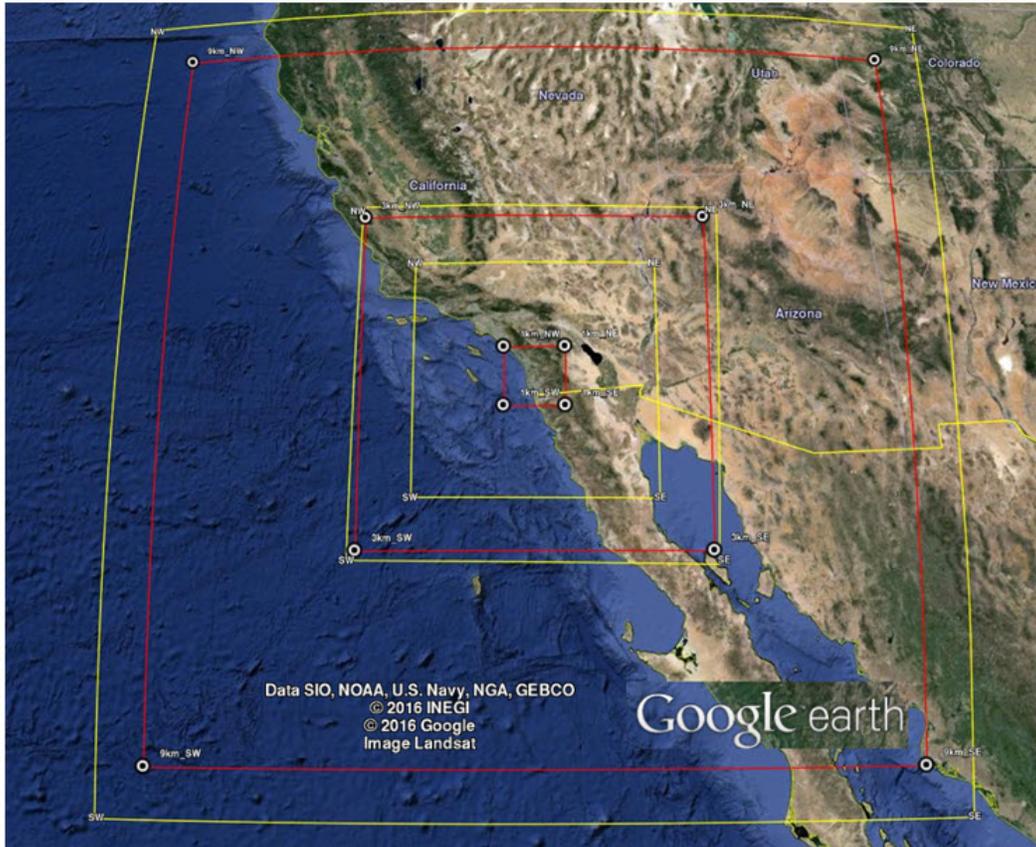


Fig. 1 Two triple-nested model domains; each domain’s center is coincident and both nested configurations are centered near San Diego, California (Google Earth 2016)

2.1 Observations

The initial conditions were constructed by starting with the GFS data as the first guess for an analysis using observations. Most observations were obtained from the Meteorological Assimilation Data Ingest System (MADIS) at the National Oceanic and Atmospheric Administration (NOAA 2014), except for the Tropospheric Airborne Meteorological Data Reporting (TAMDAR; Daniels et al. 2016) observations, which were obtained from AirDat, LLC. The MADIS database included standard surface observations, mesonet* surface observations, maritime surface observations, wind-profiler measurements, rawinsonde soundings, and Aircraft Communications Addressing and Reporting System (ACARS) data. Use and reject lists were obtained from developers of the Real-Time Mesoscale Analysis (RTMA) system (De Pondeca et al. 2011), and these were used to filter MADIS mesonet observations. This quality-assurance evaluation is especially important given the greater tendency of mesonet observations to be poorly sited compared with other, more standard, surface observations.

The Obsgrid component of WRF was used for quality control of all observations. This included gross-error checks, comparison of observations to a background field (here GFS), and comparison of observations to nearby observations. We modified Obsgrid to allow for single-level observations such as the TAMDAR and ACARS data to be more effectively compared with the GFS background field. The quality-controlled observations were output in hourly “little_r”-formatted text files for use as ground-truth data for model assessment. We employed observation nudging to the observations from these same sources for the period of 1200–1800 UTC, followed by 1-h ramping down of the nudging from 1800 to 1900 UTC, during which time no new observations are assimilated. The forecast period thus begins at 18 UTC because no observations after this time are assimilated.

2.2 Parameterizations

For the parameterization of turbulence in WRE-N, a modified version of the Mellor–Yamada–Janjić (MYJ) planetary boundary layer (Janjić 1994) scheme was used. This modification decreases the background turbulent kinetic energy and alters the diagnosis of the boundary-layer depth used for model output and data assimilation (Reen et al. 2014). The WRF single-moment, 5-class microphysics parameterization is used on all domains (Hong et al. 2004), while the Kain–Fritsch (Kain 2004) cumulus parameterization is used only on the 9-km and 15.75-km outer domains. For radiation, the Rapid Radiative Transfer Model (RRTM) parameterization (Mlawer et al. 1997) is used for long-wave radiation and the

* A network of automated meteorological observation stations.

Dudhia (1989) scheme for shortwave radiation. The Noah land-surface model (Chen and Dudhia 2001a, 2001b) is used. Additional references and other details for these parameterization schemes are available from Skamarock et al. (2008). Table 2 lists the WRE–N configurations.

Table 2 WRE-N configurations

Configuration	A–Reen (Y/N?)	B–Passner (Y/N?)	C–Dumais (Y/N?)
WRF–ARW V3.4.1	Yes	Yes	Yes
Obs-nudging FDDA	Yes	Yes	Yes
Multinest (9/3/1km)	Yes	Yes	No
Multinest (15.75/5.25/1.75km)	No	No	Yes
Cold start IC/LBC/Obs–adjusted GFS	Yes	Yes	No
MADIS observations (FDDA)	Yes	Yes	Yes
TAMDAR observations (FDDA)	Yes	Yes	Yes
Ship/buoy observations (FDDA)	Yes	Yes	Yes
Filter obs (use/reject) (FDDA)	Yes	Yes	Yes
RUNWPSPLUS QC (FDDA)	Yes	Yes	Yes
Obs-nudge rad 180,90,45	Yes	No	No
Obs-nudge rad 90,45,20	No	Yes	No
Obs-nudge rad 120,60,20	No	No	Yes
MYJ–PBL Scheme (modified)	Yes	Yes	Yes
WRF, sgl-moment, 5-class mp	Yes	Yes	Yes
Option 8—microphysics	No	Yes	Yes
Option 4—microphysics	Yes	No	No
End FDDA 360 min	Yes	Yes	Yes
End FDDA 420 min	No	No	No
Kain–Fritsch Cum Param (outer dom)	Yes	Yes	Yes
RRTM long-wave rad (Mlawer)	Yes	Yes	Yes
Shortwave rad (Dudhia)	Yes	Yes	Yes
Noah land-surface model	Yes	Yes	Yes
Fix for nudge to low water vapor	Yes	Yes	Yes
Model Top 50hPa	Yes	No	No
Model Top 10hPa	No	Yes	Yes
Feedback on	No	Yes	Yes
Obs weighting function 4E-4	No	Yes	Yes
Obs weighting function 8E-4	Yes	No	No
57 vertical levels	Yes	Yes	Yes
18-s time step	Yes	Yes	No
48-s time step	No	No	Yes

2.3 Case-Study Days

The case-study days were selected on the basis of the prevailing synoptic weather conditions over the nested domains. Table 3 has short descriptions of these conditions is provided in Table 3.

Table 3 Synoptic conditions for the case-study days considered

Case	Dates (all 2012)	Description
1	February 07–08	Upper-level trough moved onshore, which led to widespread precipitation in the region.
2	February 09–10	Quiescent weather was in place with a 500-hPa ridge centered over central California at 1200 UTC.
3	February 16–17	An upper-level low located near the California–Arizona border with Mexico at 1200 UTC brought precipitation to that portion of the domain. This pattern moved south and east over the course of the day.
4	March 01–02	A weak shortwave trough resulted in precipitation in northern California at the beginning of the period that spread to Nevada, then moved southward and decreased in coverage.
5	March 05–06	Widespread high-level cloudiness due to weak upper-level low pressure but very limited precipitation.

3. Data Preparation Using MET

The model and observational data were preprocessed into the formats required by MET Point-Stat. The WRE-N model’s output data were converted from native NetCDF formatted files to hourly gridded binary files (GRIB), Edition 1, by the WRF Unified Post Processor, which destaggers the data onto an Arakawa-A Grid. Hourly MADIS observations files in little_r format were converted into NetCDF files using a MET utility program. We used MET Point-Stat to generate the traditional, grid-to-point continuous error statistics for surface and upper air for these meteorological variables: temperature (TMP), dew-point temperature (DPT), relative humidity (RH), u-component (U) and v-component (V) winds, and wind speed and direction. Point-Stat computes matched-pair model values for the location of each observation from the MADIS data set for each forecast hour. To select observations closest to the top of the hour, and to eliminate multiple observations for each hour, we set the duplicate handling of Point-Stat to “SINGLE”. For each observation location, MET derives the forecast value corresponding to a given location by a distance-weighted mean interpolation from the enclosing grid points. In the vertical, no interpolation is used if the forecast and observation are at the same vertical level. If the levels are different, then the forecast value is interpolated to the level of the observation. MET Point-Stat produces a single matched-pair output text file and a single continuous-error statistics text file for each model output hour considered. For this study we used the MADIS parameters available at 2 m above the ground (AGL), which were TMP, RH, and DPT. The 10-m-AGL parameters were U and V.

4. Data Analysis

4.1 Value Added to WRE–N by FDDA at 1.75-km Grid Spacing

Our first step was to examine the MET model-performance statistics for the WRE–N, which was run with and without the FDDA in order to show the value added by the FDDA. The error values were computed from all matched pairs available for each hour in the Dumais 1.75-km domain, which varies (depending on the variable) over the 18-h time domain of the model. The errors for the period 1200–1800 UTC characterize the performance of the model during the assimilation phase. Figure 2 is a time-series display of the errors in the 2-m-AGL TMP’s mean error (ME: bias) and root-mean-square error (RMSE) for the Dumais 1.75-km WRE–N for Cases 1 and 5.

2m AGL temperature

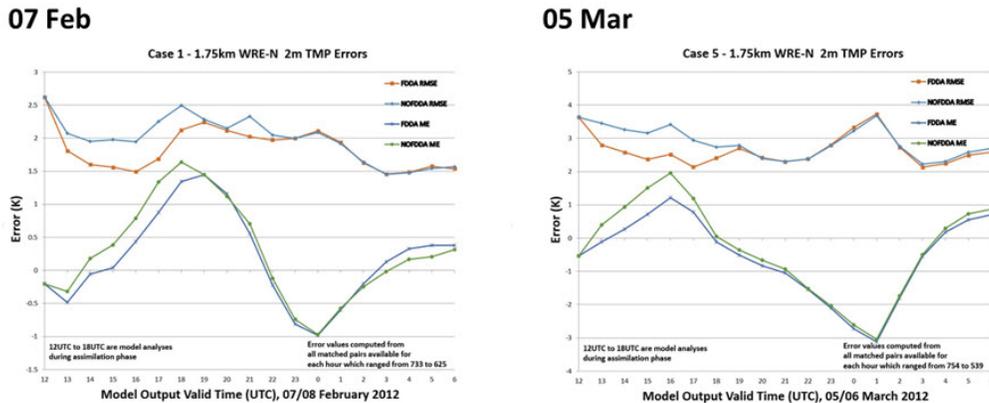


Fig. 2 Bias and RMSE errors for WRE–N with and without FDDA for 2-m-AGL TMP (in degrees Kelvin [K])

The error values were computed from all matched pairs available for each hour, which ranged from 625 to 733 over the 18-h time domain of the model for Case 1 and 539 to 754 for Case 5. There are small differences between the errors for the WRE–N with FDDA (designated as “FDDA” in the plot) and the WRE–N without FDDA (designated as “NOFDDA” in the plot) over the duration of the assimilation phase. The difference, which is nil at 1200 UTC, when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase, shows the FDDA performs slightly better than NOFDDA during the assimilation phase. Afterward, the difference is almost nil. The trends in the errors for FDDA and NOFDDA are very similar, especially during the forecast phase.

Figure 3 is a time-series display of the 2-m-AGL DPT's ME and RMSE errors for the Dumais 1.75-km WRE-N for Cases 1 and 5.

2m AGL dewpoint temperature

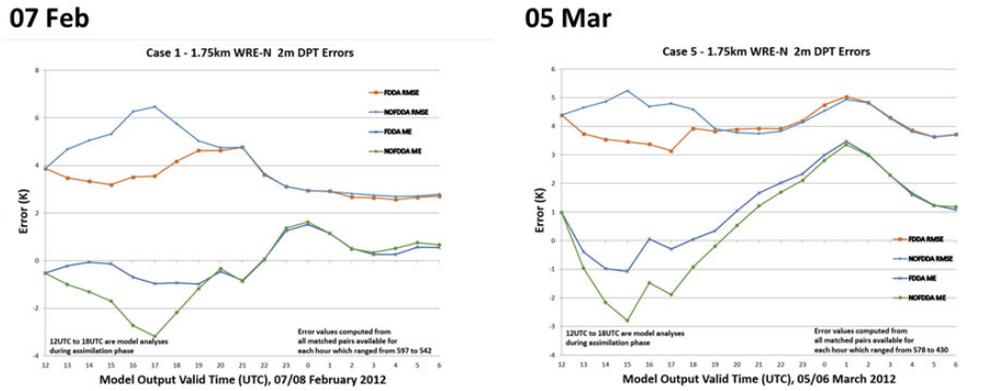


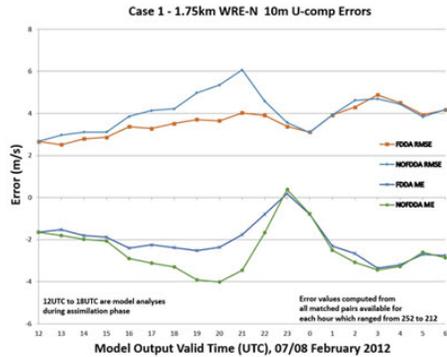
Fig. 3 Bias and RMSE errors for WRE-N with and without FDDA for 2-m-AGL DPT (K)

The error values were computed from all matched pairs available for each hour, which ranged from 542 to 597 over the 18-h time domain of the model for Case 1 and 430 to 578 for Case 5. There are larger differences between the errors for the WRE-N with FDDA and NOFDDA over the duration of the assimilation phase than were seen for TMP. The difference, which is nil at 1200 UTC, when the WRE-N is effectively the same as the GFS at the beginning of the assimilation phase, shows the FDDA performs better than NOFDDA during the assimilation phase. Afterward, the difference is almost nil. The trends in the errors for FDDA and NOFDDA are very similar, especially during the forecast phase.

Figure 4 is a time-series display of the 10-m-AGL U-component wind's ME and RMSE errors for the Dumais 1.75-km WRE-N for Cases 1 and 5.

10m AGL U wind component

07 Feb



05 Mar

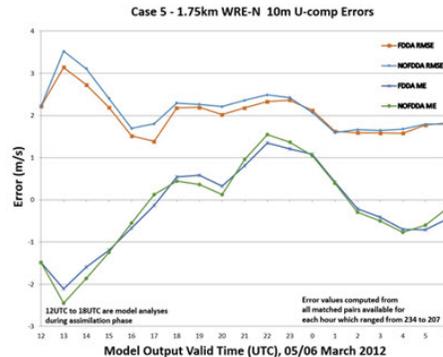


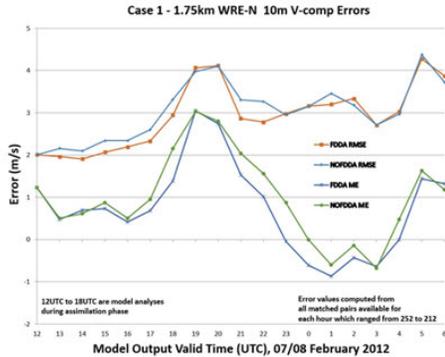
Fig. 4 Bias and RMSE errors for WRE–N with and without FDDA for 10-m-AGL U (m/s)

The error values were computed from all matched pairs available for each hour, which ranged from 212 to 252 over the 18-h time domain of the model for Case 1 and 207 to 234 for Case 5. There are modest differences between the errors for the WRE–N with FDDA and NOFDDA over the duration of the assimilation phase for Case 1 but not for Case 5. For Case 1, the difference, which is nil at 1200 UTC when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase, shows the FDDA performs better than NOFDDA during the assimilation phase. Afterward, the difference is almost nil. For Case 5, the errors for FDDA and NOFDDA are effectively the same throughout the model’s time domain. The trends in the errors for FDDA and NOFDDA for both cases are very similar, especially during the forecast phase.

Figure 5 is a time-series display of the 10-m-AGL V-component wind’s ME and RMSE errors for the Dumais 1.75-km WRE-N for Cases 1 and 5.

10m AGL V wind component

07 Feb



05 Mar

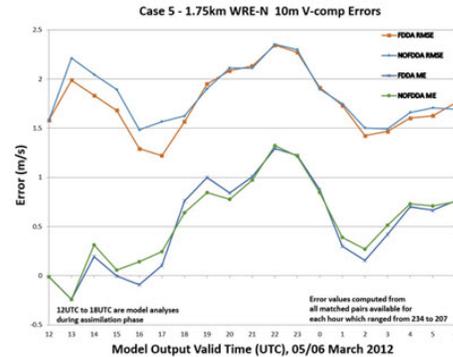


Fig. 5 Bias and RMSE errors for WRE–N with and without FDDA for 10-m-AGL V (m/s)

The error values were computed from all matched pairs available for each hour, which ranged from 212 to 252 over the 18-h time domain of the model for Case 1 and 207 to 234 for Case 5. There are small differences between the errors for the WRE–N with FDDA and NOFDDA over the duration of the assimilation phase for both cases. For Case 1, the differences—nil at 1200 UTC, when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase—show the FDDA performs slightly better than NOFDDA, but these differences do not decrease markedly following the assimilation phase as was the case for the other variables. On the other hand, for Case 5 the FDDA and NOFDDA differences are again small but decrease following the assimilation phase, which has been the case for the other variables. Afterward, the difference is almost nil. The trends in the errors for FDDA and NOFDDA for both cases are very similar, especially during the forecast phase.

4.2 Value Added to WRE–N by FDDA at 5.25- and 1.75-km Grid Spacing

The next step of our analysis was to examine the differences between FDDA and NOFDDA for the Dumais WRE–N over the 5.25-km domain and the 1.75-km domain for the same case, which was Case 4, to see any differences that are attributable to grid spacing.

Figure 6 is a time-series display of the 2-m-AGL TMP’s ME and RMSE errors for the Dumais 5.25- and 1.75-km WRE–N for Case 4.

2m AGL temperature

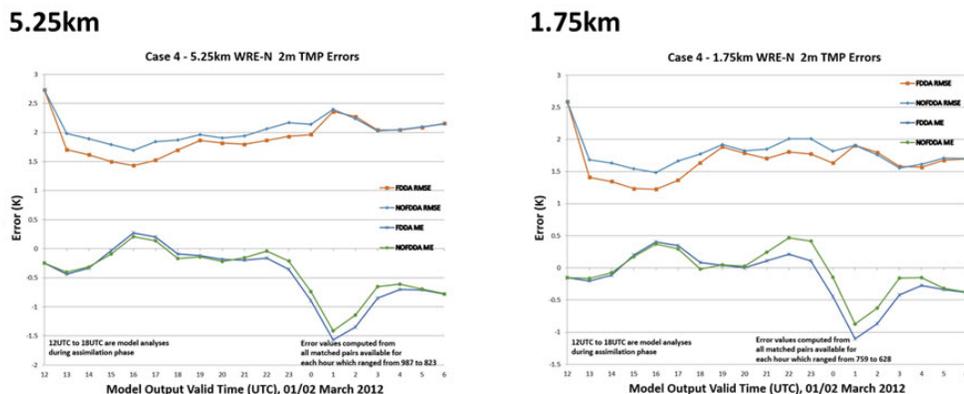


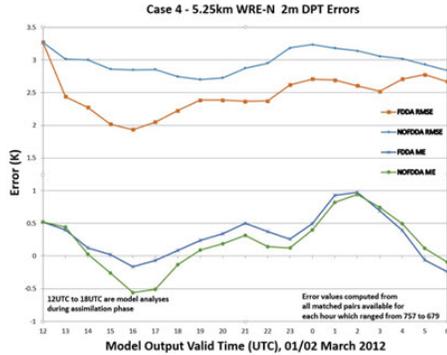
Fig. 6 Bias and RMSE errors for WRE–N with and without FDDA for 2-m-AGL TMP (K)

The error values were computed from all matched pairs available for each hour, which ranged from 823 to 987 over the 18-h time domain of the model for the 5.25-km domain and 628 to 759 for the 1.75-km domain. There are small differences between the errors for the WRE–N with FDDA and NOFDDA over the duration of the assimilation phase for both domains. For both domains, the differences—nil at 1200 UTC, when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase—show the FDDA performs slightly better than NOFDDA in terms of the RMSE, but these differences do not decrease markedly following the assimilation phase as was the case for the other variables. The difference in RMSE decreases much later in the forecast phase. For the bias, the FDDA and NOFDDA differences are very small with no clear signal that FDDA was better than the NOFDDA. The trends in the errors for FDDA and NOFDDA for both domains are very similar throughout the entire time domain with no significant differences attributable to grid spacing.

Figure 7 is a time-series display of the 2-m-AGL DPT’s ME RMSE errors for the Dumais 5.25- and 1.75-km WRE–N for Case 4.

2m AGL dewpoint temperature

5.25km



1.75km

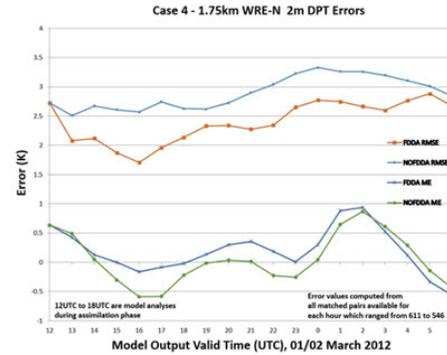


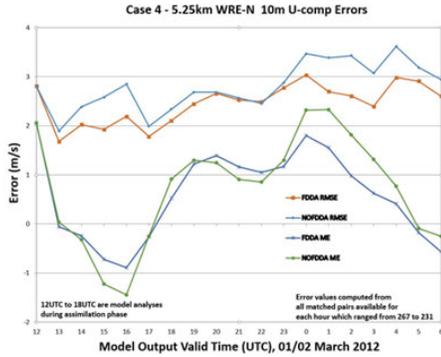
Fig. 7 Bias and RMSE errors for WRE–N with and without FDDA for 2-m-AGL DPT (K)

The error values were computed from all matched pairs available for each hour, which ranged from 679 to 757 over the 18-h time domain of the model for the 5.25-km domain and 546 to 611 for the 1.75-km domain. There are small differences between the bias errors for the WRE–N with FDDA and NOFDDA over most of the model’s time domain for both domains. However, the difference for RMSE for both domains is larger than any seen so far in the analysis. For both domains, the differences in RMSE—nil at 1200 UTC, when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase—show the FDDA performs better than NOFDDA for most of the model run. For the bias, the FDDA and NOFDDA differences are very small with no clear signal that FDDA was better than the NOFDDA. The trends in the errors for FDDA and NOFDDA for both domains are very similar throughout the entire time domain with no significant differences attributable to grid spacing.

Figure 8 is a time-series display of the 10-m-AGL U’s ME and RMSE errors for the Dumais 5.25- and 1.75-km WRE–N for Case 4.

10m AGL U wind component

5.25km



1.75km

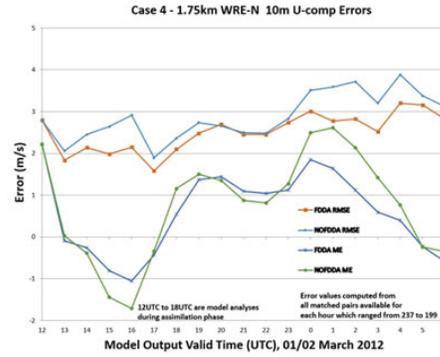


Fig. 8 Bias and RMSE errors for WRE–N with and without FDDA for 10-m-AGL U (m/s)

The error values were computed from all matched pairs available for each hour, which ranged from 231 to 267 over the 18-h time domain of the model for the 5.25-km domain and 199 to 237 for the 1.75-km domain. The differences between the bias and RMSE errors for the WRE–N with FDDA and NOFDDA vary between nil to moderate over the entire model’s time domain for both domains. The pattern of larger differences being confined only to the assimilation phase observed previously is not the case here. For both domains, the differences in RMSE—nil at 1200 UTC, when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase—show the FDDA performs better than NOFDDA for most of the model’s run. For the bias, the FDDA and NOFDDA differences are smaller with no clear signal that FDDA was better than the NOFDDA. The trends in the errors for FDDA and NOFDDA for both domains are very similar throughout the entire time domain with no significant differences attributable to grid spacing.

Figure 9 is a time-series display of the 10-m-AGL V’s ME and RMSE errors for the Dumais 5.25- and 1.75-km WRE–N for Case 4.

10m AGL V wind component

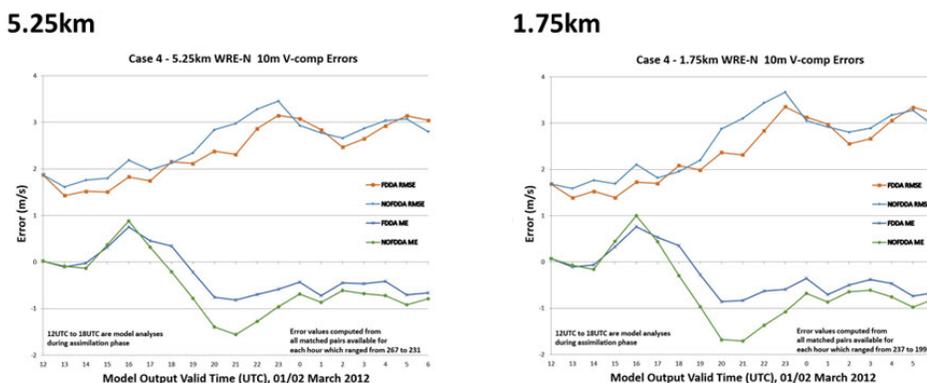


Fig. 9 Bias and RMSE errors for WRE–N with and without FDDA for 10-m-AGL V (m/s)

The error values were computed from all matched pairs available for each hour, which ranged from 231 to 267 over the 18-h time domain of the model for the 5.25-km domain and 199 to 237 for the 1.75-km domain. The differences between the bias and RMSE for the WRE–N with FDDA and NOFDDA vary between nil to moderate over the model’s time domain for both domains. The largest differences for both errors begin near the end of the assimilation phase and end at approximately 2300 UTC. For both domains, the differences in RMSE—nil at 1200 UTC, when the WRE–N is effectively the same as the GFS at the beginning of the assimilation phase—show the FDDA performs better than NOFDDA for most of the model’s run. For the bias, the FDDA and NOFDDA differences are smaller with no clear signal that FDDA was better than the NOFDDA through 1700 UTC; but, afterwards, there appears to be a clearer signal that the FDDA performed better. The trends in the errors for FDDA and NOFDDA for both domains are very similar throughout the entire time domain with no significant differences attributable to grid spacing.

4.3 Value Added by Decreased Grid Spacing

The next step of our analysis was to examine the error differences between larger and smaller grid spacing for the WRE–N with FDDA. This was accomplished by scoring each of the 3 grids using the same observations present in the innermost domain for the runs produced by Dumais and Passner for the same case, which was Case 1, to see any differences that are attributable to grid spacing.

Figure 10 is a time-series display of the 2-m-AGL TMP's ME and RMSE errors for the Dumais and Passner runs for Case 1.

07 FEB 2012 Case Study – 2m temperature

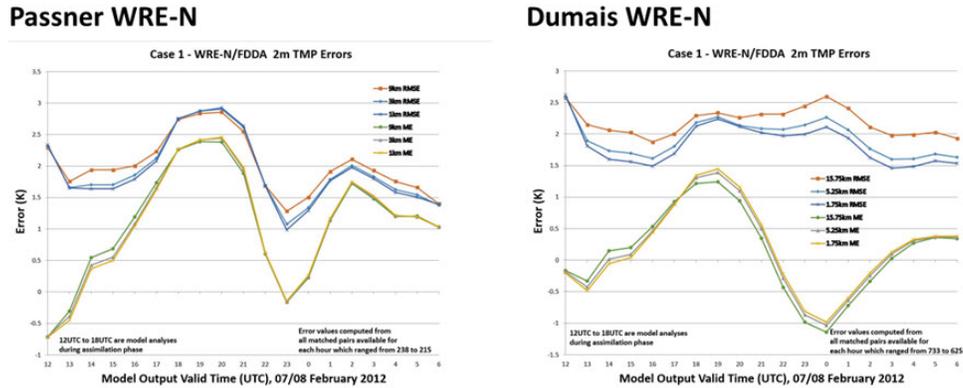


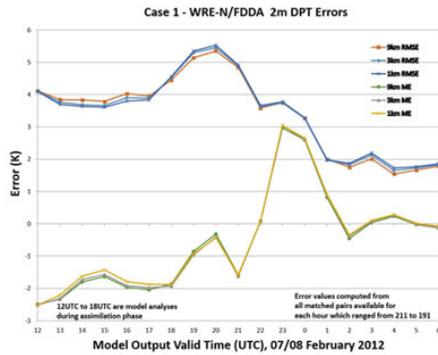
Fig. 10 Bias and RMSE errors for the 3 grids for Dumais and Passner WRE–N with FDDA for 2-m-AGL TMP (K)

The error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 215 to 238 over the 18-h time domain of the model for the Passner WRE–N and 625 to 733 for the Dumais WRE–N. For the Passner WRE–N the differences among the 3 grids are very small to nil for both bias and RMSE over the model's entire time domain. There is a distinct sinusoidal trend in bias from near zero to large positive values, which is reflected in the trend of the RMSE errors. The first maximum value for both errors occurs upon completion of the assimilation phase, which is followed by a trend toward lower-magnitude errors between 2100 and 2300 UTC, finally returning to larger errors by 0000 UTC until the end of the run. For the Dumais WRE–N, the differences among the grids are also very small to nil for bias, but not for RMSE. For RMSE—except for the beginning of the assimilation phase—there are differences in the error magnitude among the 3 grids with the smallest spacing having the smallest error and the largest spacing having the largest error, which is the expected relationship. The sinusoidal trend in RMSE seen on the Passner errors is somewhat reflected in the Dumais RMSE errors, except the amplitude for the RMSE errors is considerably reduced with the second minimum absent even though it is present in the trend of the bias errors.

Figure 11 is a time-series display of the 2-m-AGL DPT's ME and RMSE errors for the Dumais and Passner runs for Case 1.

07 FEB 2012 Case Study – 2m dewpoint

Passner WRE-N



Dumais WRE-N

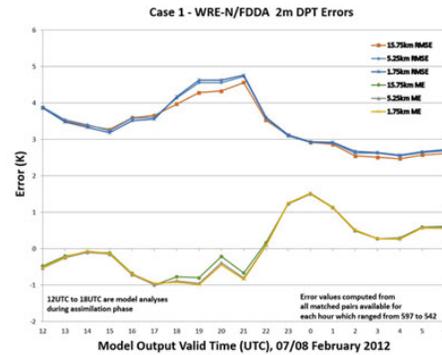


Fig. 11 Bias and RMSE errors for the 3 grids for Dumais and Passner WRE–N with FDDA for 2-m-AGL DPT (K)

The error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 191 to 211 over the 18-h time domain of the model for the Passner WRE–N and 542 to 597 for the Dumais WRE–N. For the Passner WRE–N the differences among the 3 grids for DPT are very small to nil for both bias and RMSE over the model’s entire time domain, which is the same as was seen for TMP. There is a less distinct sinusoidal trend in bias that is somewhat centered on zero with variations above and below. The same pattern is not matched in the trend of the RMSE errors. The first maximum value for RMSE occurs upon completion of the assimilation phase, which is followed by a trend toward lower-magnitude errors between 2100 and 2300 UTC, then continues a decreasing trend over the remainder of the period. For bias, the first maximum occurs a few hours later, followed by a return to near zero bias for the remainder of the period. For the Dumais WRE–N, the differences among the grids are also very small to nil for bias and RMSE. For bias and RMSE, there is also a sinusoidal pattern with the first maximum for RMSE occurring a few hours earlier than that for bias, similar to the Passner errors.

Figure 12 is a time-series display of the 10-m-AGL U’s ME and RMSE errors for the Dumais and Passner runs for Case 1.

07 FEB 2012 Case Study – 10m U wind

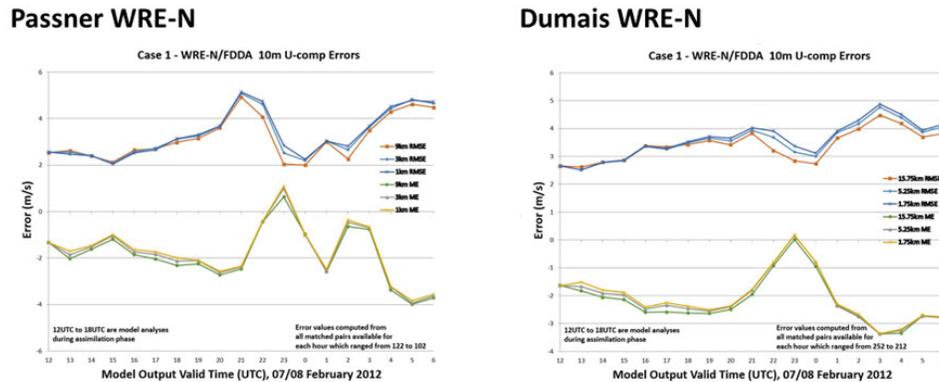


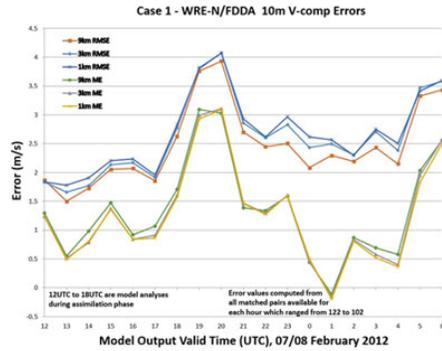
Fig. 12 Bias and RMSE errors for the 3 grids for Dumais and Passner WRE–N with FDDA for 10-m-AGL U (m/s)

The error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 102 to 122 over the 18-h time domain of the model for the Passner WRE–N and 212 to 252 for the Dumais WRE–N. For the Passner WRE–N the differences among the 3 grids for U are very small to nil for both bias and RMSE over the model’s entire time domain, which is the same as was seen for TMP. There is a sinusoidal trend in bias that is somewhat centered on -1.0 with variations above and below. The first maxima for RMSE and bias do not coincide, with the one for RMSE occurring about 2 hours before that for bias. The first maxima values for bias and RMSE occur a few hours after the completion of the assimilation phase. For the Dumais WRE–N, the differences among the grids are also very small to nil for bias and RMSE. For bias and RMSE, there is also a sinusoidal pattern with the first maximum for RMSE occurring 2 hours earlier than that for bias, similar to the Passner errors.

Figure 13 is a time-series display of the 10-m-AGL V’s ME and RMSE errors for the Dumais and Passner runs for Case 1.

07 FEB 2012 Case Study – 10m V wind

Passner WRE-N



Dumais WRE-N

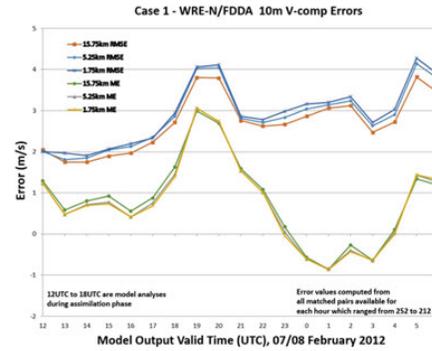


Fig. 13 Bias and RMSE errors for the 3 grids for Dumais and Passner WRE–N with FDDA for 10-m-AGL V (m/s)

The error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 102 to 122 over the 18-h time domain of the model for the Passner WRE–N and 212 to 252 for the Dumais WRE–N. For the Passner WRE–N the differences among the 3 grids for U are very small to nil for both bias and RMSE over the assimilation phase and a few hours into the forecast period. Afterward, there appears to be more separation in the errors among the 3 grids, with the lowest value being that of the 9-km grid and the highest being that of the 1-km grid, which is the opposite of the expected relationship. There is a sinusoidal trend in RMSE and bias, with the first maxima occurring just after the completion of the assimilation phase. For the Dumais WRE–N, the differences among the grids are also very small to nil for bias and RMSE over the model's entire time domain. For bias and RMSE, there is also a sinusoidal pattern, with the first maximum for RMSE and bias coincident in the hours following the assimilation phase, similar to the pattern of the Passner errors.

4.4 Performance Comparison of 2 WRE–N Configurations

The next step of our analysis was to compare the performance of 2 different configurations of the 1-km WRE–N that were run over the same domain. The 2 configurations used were those run by Passner and Reen, with FDDA in both cases. Table 4 describes the differences in configuration settings of the 2 models.

Table 4 WRE–N configuration differences

Setting/configuration	Reen	Passner
Obs-nudging radius	180, 90, 45	90, 45, 20
Microphysics option	4	8
Obs-weighting function	8E-4	4E-4
Model top	50hPa	10hPa
Feedback on (Y/N?)	No	Yes

Figure 14 is a time-series display of the 2-m-AGL TMP’s ME and RMSE errors for the Reen and Passner 1-km runs for Cases 1 and 2.

07 FEB, 09 FEB 2012 Case Studies 2m AGL temperature

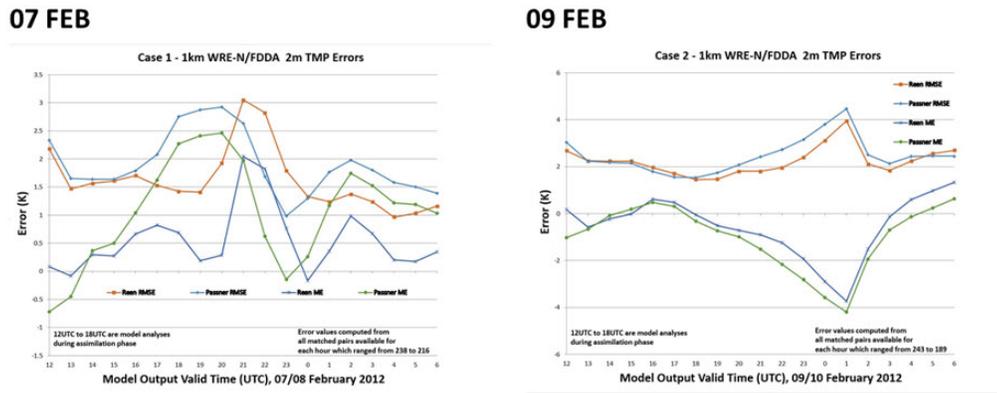


Fig. 14 Bias and RMSE errors for the Reen and Passner 1-km WRE–N with FDDA for 2-m-AGL TMP (K)

The TMP error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 216 to 238 over the 18-h time domain of the models for Case 1 and 189 to 243 for Case 2. For Case 1, both models have different error values over most of the model’s time domain, with the exception of the first few hours of the assimilation phase. They both show trends that have similarities at certain times. At other times, the events that occur in the trend appear displaced in time between one model and the other. An example of a displaced event is the first maximum in RMSE, which occurs between 1800 and 2100 UTC for the Passner WRE–N. The Reen WRE–N appears to have the same event, which occurs between 2000 and 2300 UTC. This event also occurs in the bias trend for both models with a similar displacement. Both models also show a second maximum in the trend of both errors at 0200 UTC. For Case 2, the values of the errors for both models are closer and the trends are very similar with no significant displacement. Each model has periods when its errors are of less

magnitude than the other model's, but there does not appear to be a consistent pattern with one model having smaller errors than the other. Without the benefit of additional information about the specific temporal changes in the atmospheric conditions over the domain, it is difficult to associate such changes with the error trends. However, the conditions in Case 1 involved more synoptic forcing from a frontal system and active precipitation with attendant upper-level dynamics than those in Case 2, which had quiescent conditions that may have contributed to a more sinusoidal pattern than in the trends for Case 1.

Figure 15 is a time-series display of the 2-m-AGL DPT's ME and RMSE errors for the Reen and Passner 1-km runs for Cases 1 and 2.

07 FEB, 09 FEB 2012 Case Studies 2m AGL dewpoint

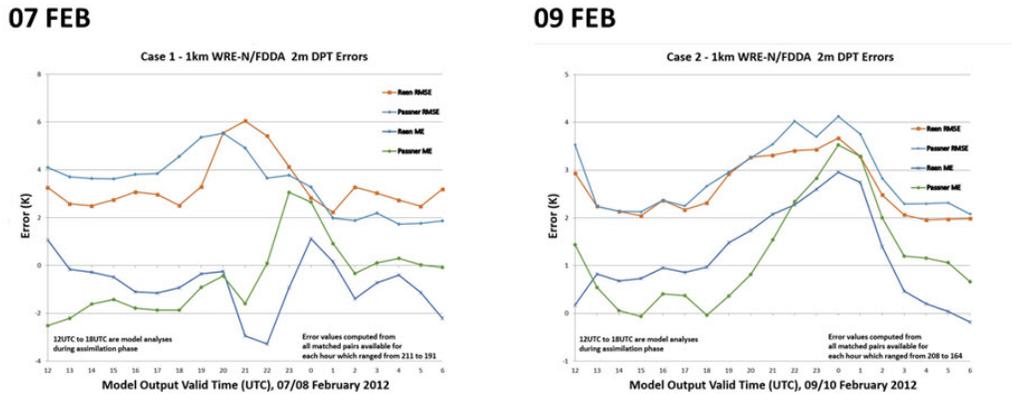


Fig. 15 Bias and RMSE errors for the Reen and Passner 1-km WRE–N with FDDA for 2-m-AGL DPT (K)

The DPT error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 191 to 211 over the 18-h time domain of the models for Case 1 and 164 to 208 for Case 2. For Case 1, both models show different error values over the course of the 18-h time domain with occurrences of a maximum in the magnitude of both errors, similar to that observed in the TMP error trend, following the assimilation phase. There is displacement of the maxima for RMSE and bias between both models, similar to that observed for TMP, with the Passner model's maximum occurring before that of the Reen model. Each model has periods when its errors are of less magnitude than the other model's but there does not appear to be a consistent pattern with one model having smaller errors than the other. For Case 2, the values of the errors for both models are similar to those of Case 1 and the trends are very similar with no significant displacement

between the maxima of the 2 models, as was the case for TMP. Each model has periods when its errors are of less magnitude than the other model's but there does not appear to be a consistent pattern with one model having smaller errors than the other.

Figure 16 is a time-series display of the 10-m-AGL U's ME and RMSE errors for the Reen and Passner 1-km runs for Cases 1 and 2.

07 FEB, 09 FEB 2012 Case Studies 10m AGL U wind

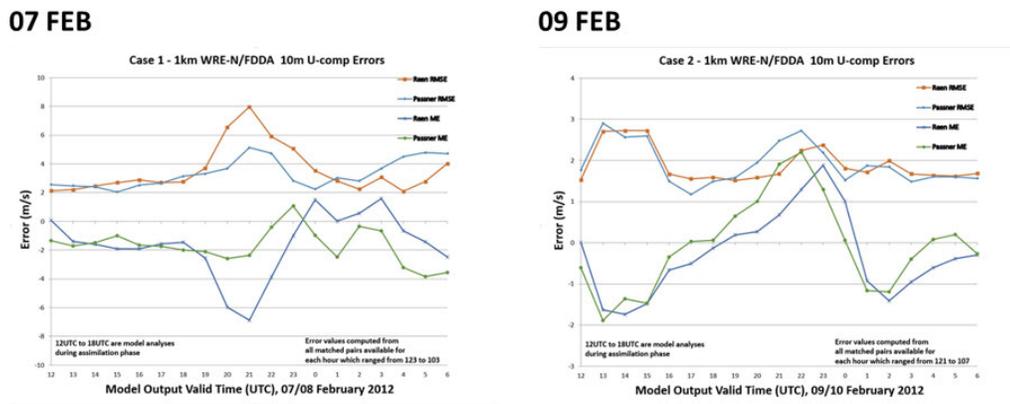


Fig. 16 Bias and RMSE errors for the Reen and Passner 1-km WRE-N with FDDA for 10-m-AGL U (m/s)

The U error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 103 to 123 over the 18-h time domain of the models for Case 1 and 107 to 121 for Case 2. For Case 1, both models show different error values over the course of the 18-h time domain with occurrences of a maximum in the magnitude both errors, similar to that observed in the TMP and DPT error trends, following the assimilation phase. It is noteworthy that, unlike for TMP and DPT, the U errors of both models are nearly identical during the assimilation phase. Regarding the maximum following the assimilation phase, there is no displacement of the maxima (in absolute value for the bias) between both models. For Case 2, the values of the errors for both models are closer and the trends are very similar with no significant displacement. There is a noteworthy shift of the timing of the postassimilation phase's maximum, placing it at approximately 2300 to 0000 UTC relative to the timing of the same event for TMP and DPT. Each model has periods when its errors are of less magnitude than the other model's, but there does not appear to be a consistent pattern over both cases with one model having smaller errors than the other.

Figure 17 is a time-series display of the 10-m-AGL V's ME and RMSE errors for the Reen and Passner 1-km runs for Cases 1 and 2.

07 FEB, 09 FEB 2012 Case Studies 10m AGL V wind

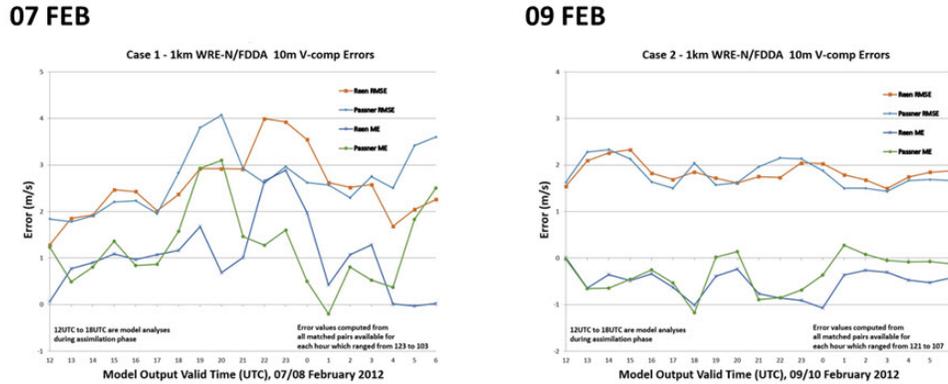


Fig. 17 Bias and RMSE errors for the Reen and Passner 1-km WRE–N with FDDA for 10-m-AGL V (m/s)

The V error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 103 to 123 over the 18-h time domain of the models for Case 1 and 107 to 121 for Case 2. For Case 1, both models show different error values over the course of the 18-h time domain with occurrences of a maximum in the magnitude both errors, similar to that observed in the TMP and DPT error trends, following the assimilation phase. It is noteworthy that, unlike for TMP and DPT, the V errors of both models are nearly identical during the assimilation phase. Regarding the maximum following the assimilation phase, unlike U, there is significant displacement of the maxima between both models. The maximum for the Passner model occurs approximately 2 h earlier than that of the Reen model. For Case 2, values of the errors for both models are significantly closer and the trends are very similar with no apparent maximum event following the assimilation phase. Each model has periods when its errors are of less magnitude than the other model's, but there does not appear to be a consistent pattern over both cases with one model having smaller errors than the other.

4.5 Performance Comparison: Dumais WRE–N with FDDA vs. the GFS

The next step of our analysis was to evaluate the performance of the Dumais WRE–N with FDDA against that of the GFS model using the results from Cases 2 and 4. The analysis could only be performed at 3-h intervals because the output of the GFS is every 3 h. In addition, the GFS does not provide the DPT variable, so RH was used instead. In this case, the scoring was accomplished over only the observations in the innermost domain for each of the 4 grids, as follows: 55-km—GFS; 15.75-km—WRE–N; 5.25-km—WRE–N; and 1.75-km—WRE–N.

Figure 18 is a time-series display of the 2-m-AGL TMP’s ME and RMSE errors for the Dumais WRE–N with FDDA and the GFS for Cases 2 and 4.

01 MAR, 09 FEB 2012 Case Studies 2m AGL temperature

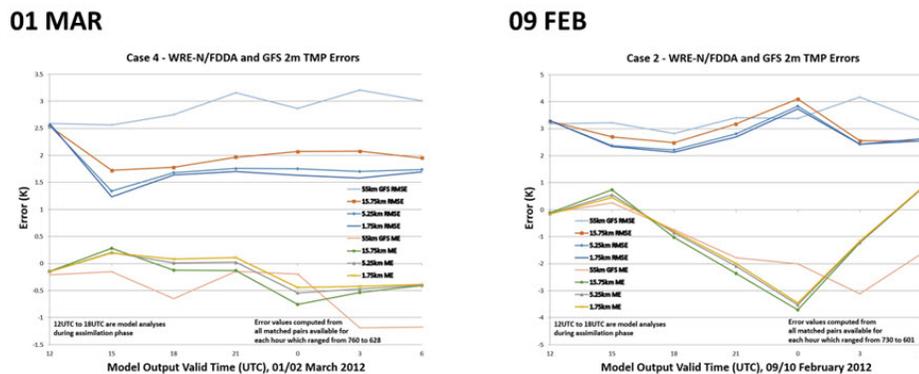


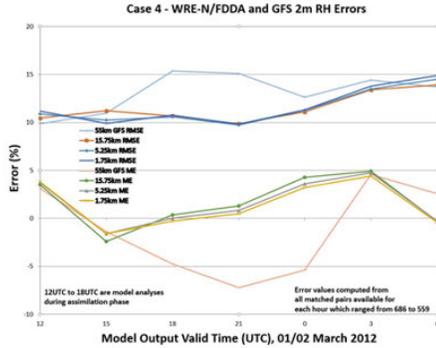
Fig. 18 Bias and RMSE errors for the Dumais WRE–N with FDDA and the GFS for 2-m-AGL TMP (K)

The TMP error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 601 to 730 over the 18-h time domain of the models for Case 2 and 628 to 760 for Case 4. For Case 4, the TMP errors at the beginning of the assimilation phase are identical because, on initialization, the GFS model supplies the background field that is interpolated onto the WRE–N grid for the first model’s lead time (Hour 0). The trend in both errors shows the WRE–N’s performance is superior at nearly all hours except for the bias at 0000 UTC. For Case 2, the superior performance of the WRE–N is not as significant until late in the forecast period, at 0300 UTC and beyond. There is even a period of 2100–0000 UTC when the bias of the GFS is smaller than that of the WRE–N.

Figure 19 is a time-series display of the 2-m-AGL RH's ME and RMSE errors for the Dumais WRE-N with FDDA and the GFS for Cases 2 and 4.

01 MAR, 09 FEB 2012 Case Studies 2m AGL relative humidity

01 MAR



09 FEB

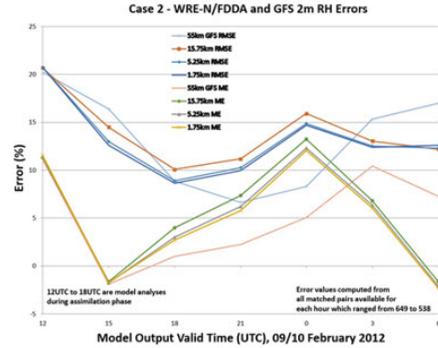


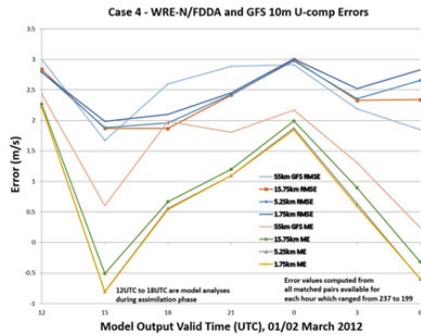
Fig. 19 Bias and RMSE errors for the Dumais WRE-N with FDDA and the GFS for 2-m-AGL RH (in %)

The RH error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 538 to 649 over the 18-h time domain of the models for Case 2 and 559 to 686 for Case 4. For Case 4, the RH errors at the beginning of the assimilation phase are all identical because, on initialization, the GFS model supplies the background field that is interpolated onto the WRE-N grid for the first model's lead time (Hour 0). The trend in both errors shows the WRE-N's performance is superior for the 3 model hours between 1800 and 0000 UTC. For Case 2, the superior performance of the WRE-N is not as apparent until late in the forecast period, at 0300 UTC and beyond. There is even a period of 1800–0000 UTC when the GFS errors are smaller than those of the WRE-N.

Figure 20 is a time-series display of the 10-m-AGL U's ME and RMSE errors for the Dumais WRE-N with FDDA and the GFS for Cases 2 and 4.

01 MAR, 09 FEB 2012 Case Studies 10m AGL U wind

01 MAR



09 FEB

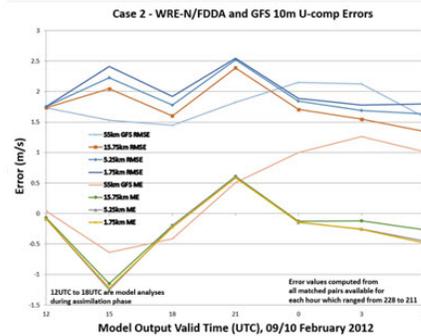


Fig. 20 Bias and RMSE errors for the Dumais WRE–N with FDDA and the GFS for 10-m-AGL U (m/s)

The U error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 211 to 228 over the 18-h time domain of the models for Case 2 and 199 to 237 for Case 4. For Case 4, the U errors at the beginning of the assimilation phase are nearly identical because, on initialization, the GFS model supplies the background field that is interpolated onto the WRE–N grid for the first model’s lead time (Hour 0). The trend in the RMSE error shows the WRE–N’s performance is superior for only 2 model hours between 1800 and 2100 UTC. For the bias, the superiority of the WRE–N is more significant over most of the 18-h period. For Case 2, the superior performance of the WRE–N is not as apparent until late in the forecast period, at 0000 UTC and beyond. For RMSE, there is even a period from 1200 to 2100 UTC when the GFS errors are smaller than those of the WRE–N.

Figure 21 is a time-series display of the 10-m-AGL V’s ME and RMSE errors for the Dumais WRE–N with FDDA and the GFS for Cases 2 and 4.

01 MAR, 09 FEB 2012 Case Studies

10m AGL V wind

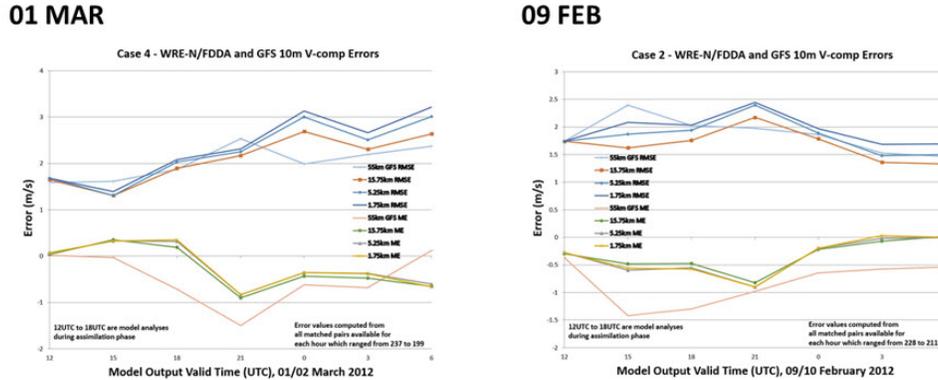


Fig. 21 Bias and RMSE errors for the Dumais WRE–N with FDDA and the GFS for 10-m-AGL V (m/s)

The V error values were computed from all matched pairs available for each hour in the innermost domain, which ranged from 211 to 228 over the 18-h time domain of the models for Case 2 and 199 to 237 for Case 4. For Case 4, the V errors at the beginning of the assimilation phase are identical because, on initialization, the GFS model supplies the background field that is interpolated onto the WRE–N grid for the first model’s lead time (Hour 0). The trend in the RMSE error shows no clear superiority of the WRE–N’s performance. After 2100 UTC, the RMSE of the GFS is slightly smaller than that of the WRE–N. For the bias, the superiority of the WRE–N is more evident in the period of 1800–0300 UTC. For Case 2, in terms of the RMSE, the superior performance of the WRE–N is not as apparent over the entire 18-h period. For bias, the WRE–N clearly performs better than the GFS for nearly the entire 18-h period.

5. Conclusions and Final Comments

From our analysis of traditional, grid-to-point, domain-level error statistics, we have found the FDDA adds value to the performance of the WRE–N—but the added value is not consistent or present for all variables, model lead times, and case-study days. From the results analyzed, the increased skill of the WRE–N at smaller grid spacing is not evident. This may be due to the difficulty in scoring, which occurs when there is a requirement for the exact match between the point observations and the forecast grid values. Often, forecasts with smaller grid spacing, despite their ability to replicate atmospheric features with more detail,

have poorer error statistics than forecasts with larger grid spacing (Ebert 2008). From the comparison of 2 different model configurations over the same domain, the results analyzed did not show the superiority of one configuration over the other. The statistics varied for both configurations depending on the lead time, the variable, and the weather situation. From the comparison of the WRE-N and the GFS, the results show that the WRE-N's performance is superior to the GFS's for varying periods of time over the course of the model's time domain, but the superiority is not consistent and there are times when the GFS scores better than the WRE-N.

The results of the analysis show strong temporal variability in the errors. This variability is characterized by maxima "events" occurring for each different WRE-N configuration at differing times and appear as displacements in time when compared on the same time-series plot. The reasons for these phase differences is not clear but may be due to the effect of temporal changes in the synoptic weather conditions occurring in the domain. More investigation is needed with numerous cases to explain the causes of these events and the influence of the weather conditions on them. In designing a subsequent assessment project to identify the underlying causes for the observed errors as well as model strengths and weaknesses, consideration must be given to the use of other verification methods.

Besides the difficulties previously described in assessing the true skill of high-resolution models, grid-to-point methods provide no information about occurrences of "near misses" that suggest a forecast of some quality or occurrences of more complete misses owing to a poor forecast. The challenge is to employ techniques that evaluate the ability of the model to replicate the features themselves, albeit with spatial displacement, in addition to the more traditional approaches. To this end, researchers have developed spatial-verification techniques that reveal more about the ability of the model to predict spatial features (Jolliffe and Stephenson 2011).

Finally, a Geographic Information System (not extensively used in atmospheric sciences) should be exploited for its ability to contextualize and analyze geospatial information—terrain type/slope, land-use effects, and other spatial and temporal variables—as explanatory metrics in model assessments (Smith et al. 2015; Smith et al. 2016a; Smith et al. 2016b). This technique has demonstrated considerable promise as an important new tool that, in addition to the traditional and non-traditional methods previously described, offers a comprehensive approach to model verification.

6. References

- Brandt J, Dawson L, Johnson J, Kirby S, Marlin D, Sauter D, Shirkey R, Swanson J, Szymber R, Zeng S. Second generation weather impacts decision aid applications and web services overview. White Sands Missile Range (NM); Army Research Laboratory (US); 2013 Jul. Report No.: ARL-TR-6525.
- Casati B, Wilson LJ, Stephenson DB, Nurmi P, Ghelli A, Pocerlich M, Damrath U, Ebert EE, Brown BG, Mason S. Forecast verification: current status and future directions. *Meteo App.* 2008;15:3–18.
- Chen F, Dudhia J. 2001a. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part II: preliminary model validation. *Mon Wea Rev.* 2001;129:587–604.
- Chen F, Dudhia J. 2001b. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity. *Mon Wea Rev.* 2001;129:569–585.
- Daniels TS, Moninger WR, Mamrosh RD. Tropospheric airborne meteorological data reporting (TAMDAR) overview. Preprints, 10th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface; 2016 Sep 1; Atlanta (GA): American Meteorological Society [accessed 2016 Aug 2].
- De Pondeca MSFV, Manikin GS, DiMego G, Benjamin SG, Parrish DF, Purser RJ, Wu WS, Horel JD, Myrick DT, Lin Y, Aune RM, Keyser D, Colman B, Mann G, Vavra J. The real-time mesoscale analysis at NOAA's National Centers for Environmental Prediction: current status and development. *Wea Forec.* 2011;26:593–612.
- Deng A, Stauffer D, Gaudet B, Dudhia J, Hacker J, Bruyere C, Wu W, Vandenberghe F, Liu Y, Bourgeois A. Update on the WRF-ARW end-to-end multi-scale FDDA system. Paper presented at: 10th WRF Users' Workshop, National Center for Atmospheric Research; Paper No.: 1.9. 2009 Jun 23–26; Boulder (CO).
- 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.

- Dumais R, Kirby S, Flanigan R. Implementation of the WRF four-dimensional data assimilation method of observaion nudging for use as an ARL weather running estimate-nowcast. White Sands Missile Range (NM): Army Research Laboratory (US); 2013 Jun. Report No.: ARL-TR-6485.
- Dumais RE, Reen BP. Data assimilation techniques for rapidly relocatable weather research and forecasting modeling. White Sands Missile Range (NM): Army Research Laboratory (US); 2013 Jun. Report No.: ARL-TN-0546.
- Dumais RE Jr, Henmi T, Passner J, Jameson T, Haines P, Knapp D. A mesoscale modeling system developed for the U.S. Army. White Sands Missile Range (NM): Army Research Laboratory (US); 2004 Apr. Report No.: ARL-TR-3183.
- Ebert E, Wilson L, Weigel A, Mittermaier M, Nurmi P, Gill P, Göber M, Joslyn S, Brown B, Fowler T, Watkins A. Progress and challenges in forecast verification. *Meteo App.* 2013;20(2):130–139.
- Ebert E. Fuzzy verification of high resolution gridded forecasts: a review and proposed framework. *Meteo App.* 2008;15:51–64.
- [EMC] Environmental Modeling Center. The GFS atmospheric model. NCEP Office Note No.: 442; 2003.
- Google Earth. Mountain View (CA); 2016 [accessed 2016 Aug 24]. http://maps.google.com/help/terms_maps.html.
- Hong SY, Dudhia J, Chen S-H. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon Wea Rev.* 2004;132:103–120.
- Janjić ZI. The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon Wea Rev.* 1994;122:927–945.
- Jolliffe IT, Stephenson DB, editors. *Forecast verification: a practitioner's guide in atmospheric science*. 2nd ed. Hoboken (NJ): John Wiley and Sons; 2011.
- Kain JS. The Kain-Fritsch convective parameterization: an update. *J App Meteo.* 2004;43:170–181.
- Liu Y, Bourgeois A, Warner T, Swerdlin S, Hacker J. Implementation of observation-nudging based FDDA into WRF for supporting ATEC test operations. Proceedings of the 6th WRF/15th MM5 Users' Workshop, National Center for Atmospheric Research; 2005 Jun 27; Boulder (CO).

- Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J Geoph Res Atmos.* 1997;102:16663–16682.
- [NCAR] National Center for Atmospheric Research. Model evaluation tools version 4.1 (METv4.1), user's guide 4.1. Boulder, CO; 2013 May.
- [NOAA] Meteorological assimilation data ingest system (MADIS). College Park (MD): National Oceanic and Atmospheric Administration. [accessed 2016 Jul 27]. <http://madis.noaa.gov>.
- [NRC] National Research Council. When weather matters: science and service to meet critical societal needs. Washington (DC): The National Academies Press; 2010.
- Raby JW. Application of a fuzzy verification technique for assessment of the Weather Running Estimate–Nowcast (WRE–N) model. White Sands Missile Range (NM): Army Research Laboratory (US); 2016 Oct. Report No.: ARL-TR-7849
- Raby JW, Cai H. Verification of spatial forecasts of continuous meteorological variables using categorical and object-based methods. White Sands Missile Range (NM): Army Research Laboratory (US); 2016 Aug. Report No.: ARL-TR-7751.
- Raby J, Passner J, Vaucher G, Raby Y. Performance comparison of high resolution weather research and forecasting model output with north american mesoscale model initialization grid forecasts. White Sands Missile Range (NM): Army Research Laboratory (US); 2012 May. Report No.: ARL-TR-6000.
- Reen BP, Schmehl KJ, Young GS, Lee JA, Haupt SE, Stauffer DR. Uncertainty in contaminant concentration fields resulting from atmospheric boundary layer depth uncertainty. *J App Meteo Clim.* 2014;53:2610–2626.
- Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X-Y, Wang W, Powers JG. A description of the advanced research WRF version 3. Boulder (CO): National Center for Atmospheric Research (US); 2008 Jun. NCAR Technical Note No.: TN-475+STR.
- Smith JA, Foley TA, Raby JW, Reen B. Investigating surface bias errors in the Weather Research and Forecasting (WRF) model using a Geographic Information System (GIS). White Sands Missile Range (NM): Army Research Laboratory (US); 2015 Feb. Report No.: ARL-TR-7212.

- Smith JA, Foley TA, Raby JW, Reen BP, Penc RS. 2016a. Case study applying GIS tools to verifying forecasts over a domain. Paper presented at: 96th annual American Meteorological Society Meeting, 23rd Conference on Probability and Statistics in the Atmospheric Sciences; 2016 Jan 10–14; New Orleans (LA). Paper No.: 13.3.
- Smith JA, Raby JW, Foley TA, Reen BP, Penc RS. 2016b. Case study applying GIS tools to verifying forecasts over a mountainous domain. Paper presented at: 17th Mountain Meteorology Conference, American Meteorological Society. 2016 Jun 27–July 1; Burlington (VT). Paper No.: 2.5.
- Stauffer DR, Seaman NL. Multiscale four-dimensional data assimilation. *J App Meteo.* 1994;33:416–434.
- Wilks DS. *Statistical methods in the atmospheric sciences*. 3rd ed. Oxford (UK): Academic Press, 2011.

List of Symbols, Abbreviations, and Acronyms

ACARS	Aircraft Communications, Addressing, and Reporting System
AGL	above ground level
ARL	US Army Research Laboratory
ARW	Advanced Research Weather Research and Forecasting model
DPT	dew-point temperature
FDDA	Four-Dimensional Data Assimilation
GFS	Global Forecast System
GRIB	gridded binary file
hPa	hectopascal
K	degrees Kelvin
MADIS	Meteorological Assimilation Data Ingest System
ME	mean error
MET	Model Evaluation Tools
MYJ	Mellor–Yamada–Janjic
MyWIDA	My Weather Impacts Decision Aid
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NOFDDA	“no FDDA”
NWP	Numerical Weather Prediction
PBL	planetary boundary layer
RH	relative humidity
RMSE	root-mean-square error
RRTM	Rapid Radiative Transfer Model
RTMA	Real-Time Mesoscale Analysis
TAMDAR	Tropospheric Airborne Meteorological Data Reporting
TMP	temperature

U	u-component wind
UTC	coordinated universal time
V	v-component wind
WRE-N	Weather Running Estimate-Nowcast
WRF	Weather Research and Forecasting
WRF-ARW	Weather Research and Forecasting, Advanced Research WRF

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

2 DIRECTOR
(PDF) US ARMY RESEARCH LAB
RDRL CIO L
IMAL HRA MAIL & RECORDS
MGMT

1 GOVT PRINTG OFC
(PDF) A MALHOTRA

12 US ARMY RSRCH LAB
(PDF) ATTN RDRL CIE M
J RABY
H CAI
G VAUCHER
D KNAPP
J SMITH
J PASSNER
R PENC
S KIRBY
R DUMAIS
T JAMESON
B REEN
B MACCALL

2 US ARMY RSRCH LAB
(PDF) ATTN RDRL CIE
P CLARK
R RANDALL

2 US ARMY RSRCH LAB
(PDF) ATTN RDRL CIE D
S O'BRIEN
J JOHNSON

1 US NAVY RSRCH LAB
(PDF) J MCLAY

1 US AIR FORCE 557TH
(PDF) WEATHER WING
R CRAIG

1 DCGS-A WEATHER EE LEAD
(PDF) J CARROLL

3 UCAR
(PDF) T FOWLER
JH GOTWAY
B BROWN

1 USAICOE
(PDF) J STALEY