ARL-TR-7751 ● AUG 2016

US Army Research Laboratory

Verification of Spatial Forecasts of Continuous Meteorological Variables Using Categorical and Object-Based Methods

by John W Raby and Huaqing Cai

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.

ARL-TR-7751 ● AUG 2016

US Army Research Laboratory

Verification of Spatial Forecasts of Continuous Meteorological Variables Using Categorical and Object-Based Methods

by John W Raby and Huaqing Cai
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
August 2016	Technical Report	October 2015–July 2016

4. TITLE AND SUBTITLE
Verification of Spatial Forecasts of Continuous Meteorological Variables Using Categorical and Object-Based Methods

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
John W Raby and Huaqing Cai

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-7751

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
Spatial forecasts from Numerical Weather Prediction (NWP) models of meteorological variables to support Army operations on the battlefield have become an integral part of the products available for the Staff Weather Officer to use in providing mission planning and execution forecasts. These forecasts are ingested by Army tactical decision aids (TDAs) and are fused with information on the operational weather thresholds, which impact the performance of Army systems and missions. The TDA generates spatial forecasts of these impacts for user-specified systems and/or missions. This report presents methods to verify forecast fields of meteorological variables that have been filtered by the application of a threshold, the same way as that used by the TDA. A threshold applied to a continuous variable field becomes a categorical forecast for which there are traditional and nontraditional methods for verification. This study evaluates the ability of the NWP model to predict a category of the spatial variable. Preliminary results suggest a combination of a traditional technique for assessing categorical forecasts with a nontraditional object-based, forecast-evaluation technique has great potential in assessing forecasts of continuous variables that have been filtered by the application of a threshold.

15. SUBJECT TERMS
object-based, forecast, verification, categorical forecast, weather impacts, thresholds, Numerical Weather Prediction, observations, Model Evaluation Tools

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			John W Raby
Unclassified	Unclassified	Unclassified	UU	40	19b. TELEPHONE NUMBER (Include area code) 575-678-2004

Contents

List of Figures

List of Tables

Acknowledgments

We offer our thanks to Mr Robert Dumais of the US Army Research Laboratory (ARL) who contributed data and information without which the study could not have been completed.

Many thanks to Mr Martin Kufus of ARL Technical Publishing at White Sands Missile Range, New Mexico, for his many helpful suggestions and consistently high standard of editing.

Executive Summary

Spatial forecasts from Numerical Weather Prediction (NWP) models of tactically significant meteorological variables to support Army operations on the battlefield have become an integral part of the products available for the Air Force Staff Weather Officer to use in providing mission planning and execution forecasts. These forecasts are ingested by Army tactical decision aids (TDAs). These TDAs fuse information on the characteristic operational weather thresholds that affect the performance of Army systems and missions with the spatial forecast information from NWP to generate spatial forecasts of these impacts for user-specified systems and/or missions for the time period and location of interest. This report presents methods that can verify spatial forecast fields of meteorological variables that have been filtered by the application of a threshold the same way as that used by the TDA. In effect, a threshold applied to a continuous variable field becomes a categorical forecast for which there are traditional and nontraditional methods for verification. This study evaluates the ability of the NWP model to predict a category of the spatial variable.

Traditional methods have been developed to verify the skill of NWP to predict categories of continuous meteorological variables. These methods apply the established theoretical framework for evaluating deterministic binary forecasts. This framework involves defining a binary event through the application of a category or threshold and evaluates the forecast skill by counting the numbers of times the event was forecast or not and observed or not in a contingency table. There are numerous statistics and skill scores that can be computed from the data collected by this method. For this study, we obtained forecasts from the Army's Weather Running Estimate–Nowcast (WRE–N), which is a version of the Advanced Research Weather Research and Forecasting Model adapted for generating short-range nowcasts and gridded observations produced by the National Oceanographic and Atmospheric Administration's Global Systems Division using the Local Analysis and Prediction System. A tool developed by the National Center for Atmospheric Research (NCAR) called MET Series-Analysis was used to generate the skill scores and statistics at every grid point and then graphical products that display the spatial distribution of the scores and statistics.

Nontraditional methods have been developed to assess the ability of NWP models to predict the occurrence of precipitation through the application of a spatial-object-based approach, which compares the attributes of the forecast areas of precipitation with those obtained from observations of precipitation areas. This method applies techniques developed for image processing and matching with the

goal of quantifying the degree to which the forecast object is analogous to an observed object. This involves the application of a threshold to the variable field to define the objects of interest. For this study, we used the object-based forecast-evaluation tool developed by NCAR called Method for Object-Based Diagnostic Evaluation (MODE). MODE was developed to evaluate model precipitation forecasts. For this study, a novel approach was taken by applying MODE to assess the ability of the WRE-N to predict objects in continuous meteorological variable fields.

Preliminary results suggest a combination of a traditional technique for assessing categorical forecasts with a nontraditional object-based, forecast-evaluation technique has great potential in assessing forecasts of continuous variables—especially when most TDAs rely on a specific threshold of a particular variable such as temperature to determine impacts on Army missions and systems.

1. 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. In this study, we have used the Advanced Research version of WRF (Skamarock et al. 2008) that we abbreviate as WRF–ARW. 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 (Deng et al. 2009; Stauffer and Seaman 1994). 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 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 My Weather Impacts Decision Aid (MyWIDA), an Army-developed decision aid used to determine atmospheric impacts on Army and Joint systems and operations (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 (2012) 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. Furthermore, the verification of WRE–N spatial fields of continuous meteorological variables that have been filtered by the application of a threshold to evaluate the applicability of such output for use in MyWIDA has not been accomplished.

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 557th Weather Wing (formerly the Air Force Weather Agency). NCAR is sponsored by the National Science

Foundation. MET Series-Analysis performs categorical verification of gridded model output against observations that have been analyzed and placed on a grid matching that of the model. MET Method for Object-Based Diagnostic Evaluation (MODE) has been used for object-based spatial verification of high-resolution forecast grids of precipitation.

ARL has employed MET MODE in prior assessments such as that of Cai and Dumais (2015). They evaluated the 3-km grid spacing High Resolution Rapid Refresh (HRRR) model to demonstrate the utility of a nontraditional object-based technique in providing additional information to improve model precipitation forecasts to complement the information provided by traditional verification techniques. In a separate study, Vaucher and Raby (2014) developed the capability to use MODE for object-based assessment of 1-km grid spacing WRE–N output of continuous meteorological variables. For this study, the only source of gridded observations available was from the National Oceanic and Atmospheric Administration (NOAA)–National Centers for Environmental Prediction (NCEP) Real-Time Mesoscale Analysis (RTMA) product (De Pondeca et al. 2011). In Vaucher and Raby (2014), the RTMA product, generated at a horizontal grid spacing of 2.5 km, was used with the WRE–N output that was remapped from a 1-km grid to a 2.5-km grid to produce the required matching grid.

MODE proved to be useful as an assessment tool for the WRE–N over an Army-scale domain, and plans were made to expand its use to perform evaluations of continuous meteorological variables generated by the WRE–N at 1.75-km grid spacing. Collaborations with NOAA's Global Systems Division (GSD) resulted in the generation of 1.75-km grids of observations of surface meteorological variables for the same domain as the WRE-N using the NOAA–GSD Local Analysis and Prediction System (LAPS).

The WRE–N was run with and without FDDA for 5 case-study days over a 1.75-km grid-spacing domain in Southern California over highly varied terrain and with a dense observational network that provided a robust data set of model output for analysis. The case-study days from February–March 2012 were picked to vary weather conditions from a strong synoptic forcing situation to a quiescent situation. (The weather conditions for each study day are described in Section 2.3.)

This study explores the utility of MET Series-Analysis and expands the utility of MODE for assessing the WRE–N at tactically significant grid spacings; also, it evaluates the accuracy of WRE–N spatial forecasts of continuous meteorological variables that have been filtered using a threshold similar to the way MyWIDA

uses WRE–N output to provide spatial distributions of forecast weather impacts to Army missions and systems.

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 a multi-nest set of domains to produce a fine inner mesh with 1.75-km grid spacing and 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 for the triple-nested domain. This global model for ARL development and testing has been the National Center for Environmental Prediction's Global Forecast System (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; Dumais et al. 2012).

Table 1 WRE–N triple-nested domain dimensions in km

East–West dimension	North–South dimension	Grid spacing
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 hourly model outputs, which were produced for each of 5 days in February and March 2012. The modeling domains are depicted in Fig. 1.

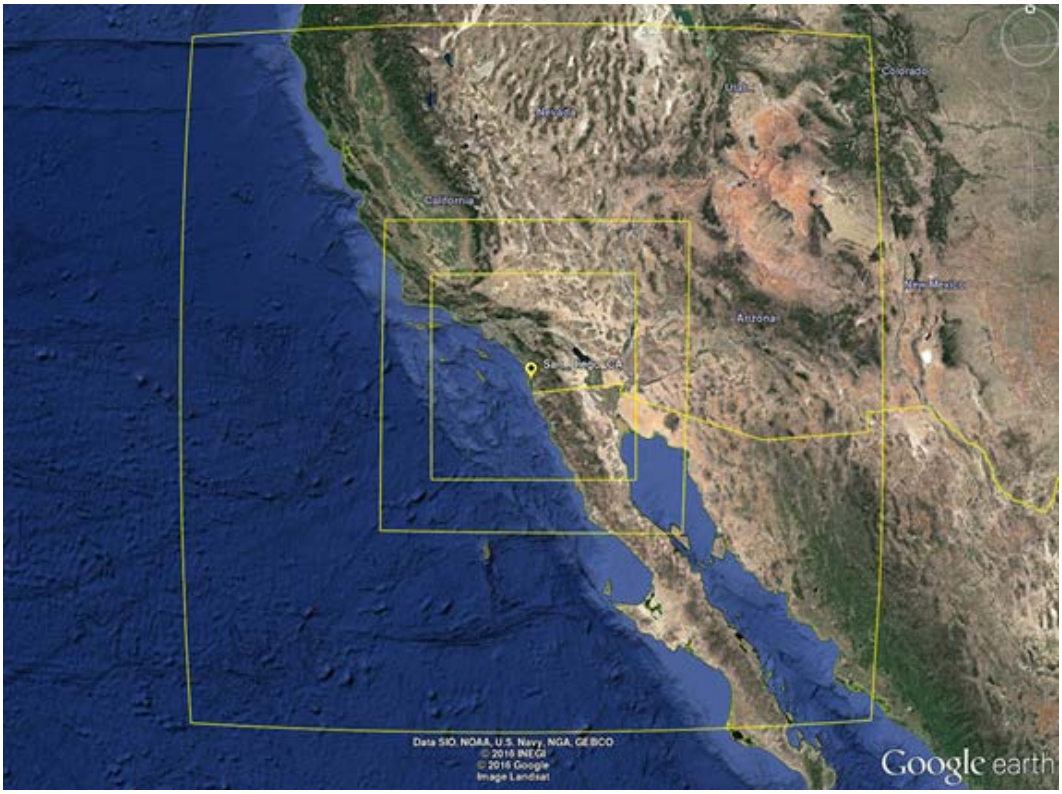

Fig. 1 Triple-nested model domains: domain center points are coincident and are centered near San Diego, California (Google Earth 2016)

2.1 Observations for Assimilation

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) (NOAA 2014), except for the Tropospheric Airborne Meteorological Data Reporting (TAMDAR) (Daniels et al. 2006) 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 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 more poorly sited than other, more standard, surface observations.

*A network of automated meteorological observation stations.

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 against 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 preforecast period of 1200–1800 UTC (0- through 6-h lead times), followed by 1 h ramping down of the nudging from 1800 to 1900 UTC, during which no new observations are assimilated. The true, free forecast period thus begins at 1800 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 (PBL) (Janjić 1994) scheme was used. This modification decreases the background turbulent kinetic energy (TKE) 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 15.75-km outer domain. For radiation, the Rapid Radiative Transfer Model (RRTM) parameterization (Mlawer et al. 1997) is used for longwave radiation and the 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 WRF configuration settings.

Table 2 WRE–N configuration

Configuration	Y/N?
WRF-ARW V3.4.1	Yes
Obs-nudging FDDA	Yes
Multinest (15.75/5.25/1.75 km)	Yes
MADIS observations (FDDA)	Yes
TAMDAR observations (FDDA)	Yes
Ship/buoy observations (FDDA)	Yes
Filter obs (use/reject) (FDDA)	Yes
RUNWPSPLUS QC (FDDA)	Yes
Obs-nudge rad 120,60,20	Yes
MYJ-PBL scheme (modified)	Yes
WRF, sgl-moment, 5-class mp	Yes
Option 8 – microphysics	Yes
End FDDA 360 mins	Yes
Kain-Fritsch cum param (outer dom)	Yes
RRTM longwave rad (Mlawer)	Yes
Short wave rad (Dudhia)	Yes
Noah land surface model	Yes
Fix for nudge to low water vapor	Yes
Model top 10 hPa	Yes
Feedback on	Yes
Obs weighting function 4E-4	Yes
57 vertical levels	Yes
48-s time step	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 provides a short description of these conditions.

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.

2.4 Observations for Verification

The LAPS gridded observation data sets produced by NOAA–GSD consisted of 12 hourly Gridded Binary format, edition 2 (GRIB2) files of 2-m above-ground-level (AGL) temperature (TMP), relative humidity (RH), and dew-point temperature (DPT), and 10-m AGL U-component and V-component winds for the period of 1200–2300 UTC (forecast lead times 0 through 11) on each of the 5 cases. The output grid used by the LAPS was 289 × 289 with 1.75-km grid spacing.

3. Data Preparation Using MET

The model and observational data were preprocessed into the formats required by MET Series-Analysis and MODE. The WRE–N model output data were converted from native Network Common Data Form (NetCDF) files to hourly Gridded Binary format, edition 1 (GRIB) files by the WRF Unified Post Processor, which destaggers the data onto an Arakawa-A Grid containing 288 × 288 grid points. The hourly GRIB2 files on a 289 × 289 grid had to be remapped to the 288 × 288 grid to match that of the WRE–N grid. The NCAR "COPYGB" utility program was used to remap the observations and convert the files to GRIB (Developmental Testbed Center 2016). We used MET Series-Analysis to generate the grid-to-grid, categorical-error statistics for surface meteorological variables TMP and DPT in degrees Kelvin (deg K), RH (%), and wind speed in meters per second (WIND). Series-Analysis computed the contingency-table statistics and skill scores for each forecast hour for 5 different thresholds (categories) at every grid point over all 12 forecast lead times and all 5 case-study days. The thresholds were specified using the FORTRAN convention of "GE" to indicate greater than or equal to the given threshold value and are shown in Table 4.

Table 4 Thresholds used in MET Series-Analysis

TMP (deg K)	DPT (deg K)	RH (%)	WIND (m/s)
270	262	25	2
275	267	40	5
280	272	55	8
285	277	70	11
290	282	85	14

MET Series-Analysis generates many categorical skill scores and contingency-table statistics. Of these, Table 5 lists those which were output initially.

Table 5 Initial Series-Analysis skill scores and contingency-table statistics

Score/statistic	Description
BASER	Base rate
FMEAN	Mean forecast value
PODY	Hit rate
FAR	False alarm ratio
FBIAS	Frequency bias
CSI	Critical success index
GSS	Gilbert skill score
ACC	Accuracy

For this study, we reduced our analysis to consider only CSI and FBIAS for the variables of 2-m AGL TMP and RH and 10-m AGL WIND to accomplish a preliminary evaluation of the utility of categorical verification in assessing the accuracy of WRE–N output that was filtered by application of a threshold. The Series-Analysis output NetCDF file was ingested into the Unidata Integrated Data Viewer, which was used to generate graphics displaying the spatial distribution of the CSI and FBIAS over the WRE–N domain (Murray et al. 2016).

We used MET MODE to generate statistics from the comparison of objects in the forecast fields and observed fields for each variable for each forecast hour for a single threshold (category) over all 12 forecast lead times and all 5 case-study days. The statistics computed were total number of objects and total area of objects defined by the threshold for each modeled and observed variable. The thresholds used were selected from those used for MET Series-Analysis and were specified as GE to the given threshold value and are shown in Table 6.

Table 6 Thresholds used in MET MODE

TMP (deg K)	DPT (deg K)	RH (%)	WIND (m/s)
290	282	85	11

4. Data Analysis

4.1 Analysis of MET Series-Analysis Results

The CSI and FBIAS are defined by a ratio of counts determined using a 2 × 2 contingency table. Table 7 shows the contingency table with notation consistent with the formulae for the scores and statistics as implemented in the MET (NCAR, 2013).

Table 7 The 2 × 2 contingency table from the MET User's Guide 4.1

2x2 contingency table in terms of counts. The n_{ij} values in the table represent the counts in each forecast-observation category, where i represents the forecast and j represents the observations. The "." symbols in the total cells represent sums across categories.

Forecast	Observation		Total
	o = 1 (e.g., "Yes")	*o* = 0 (e.g., "No")	
f = 1 (e.g., "Yes")	n_{11}	n_{10}	$n_{1.} = n_{11} + n_{10}$
f = 0 (e.g., "No")	n_{01}	n_{00}	$n_{0.} = n_{01} + n_{00}$
Total	$n_{.1} = n_{11} + n_{01}$	$n_{.0} = n_{10} + n_{00}$	$T = n_{11} + n_{10} + n_{01} + n_{00}$

The counts, n_{11}, n_{10}, n_{01}, and n_{00}, are sometimes called the "Hits", "False alarms", "Misses", and "Correct rejections", respectively.

By dividing the counts in the cells by the overall total, T, the joint proportions, p_{11}, p_{10}, p_{01}, and p_{00} can be computed. Note that $p_{11} + p_{10} + p_{01} + p_{00} = 1$. Similarly, if the counts are divided by the row (column) totals, conditional proportions, based on the forecasts (observations) can be computed.

The CSI score (Eq. 1) is computed as described in the MET User's Guide 4.1:

$$\mathrm{CSI} = \frac{n_{11}}{n_{11} + n_{10} + n_{(} \quad , \tag{1}$$

with CSI being the ratio of the number of times the event was correctly forecasted to occur to the number of times it was either forecasted or occurred. CSI ignores the "correct rejections" category (i.e., n_{00}).

The value of the CSI ranges between 0 and 1, with 1 being a perfect forecast and 0 being a forecast with no skill.

The FBIAS score is computed as described in Eq. 2:

$$\mathrm{Bias} = \frac{n_{11} + n_{10}}{n_{11} + n_{01}} = \frac{n_{1.}}{n_{.1}} \quad , \tag{2}$$

with FBIAS defined as the ratio of the total number of forecasts of an event to the total number of observations of the event. A "good" value of Frequency Bias is close to 1; a value greater than 1 indicates the event was forecasted too frequently and a value less than 1 indicates the event was not forecasted frequently enough.

4.1.1 Apply CSI for Assessment of the WRE–N

A display of the spatial distribution of the CSI for TMP is shown in Fig. 2. The CSI for TMP shows scoring in all areas with the exception of high elevations in the mountains in the central portion of the domain—as indicated by white coloring—that do not contain calculated values for CSI. This is due to the lack of event occurrences with which to calculate a score. This is not inconsistent with the expectations for lower TMPs at higher elevation. The higher values of CSI are located inland over lower-elevation areas with the exception of the Salton Sea, which has anomalously low CSI. Subsequent investigation of the LAPS gridded observations revealed that the land-surface model used was of insufficient resolution to adequately distinguish the water area of the Salton Sea from the surrounding land area; thus, it did not provide a good ground-truth representation for this area. The CSI over the ocean is homogeneously near zero while over land the CSI varies irregularly.

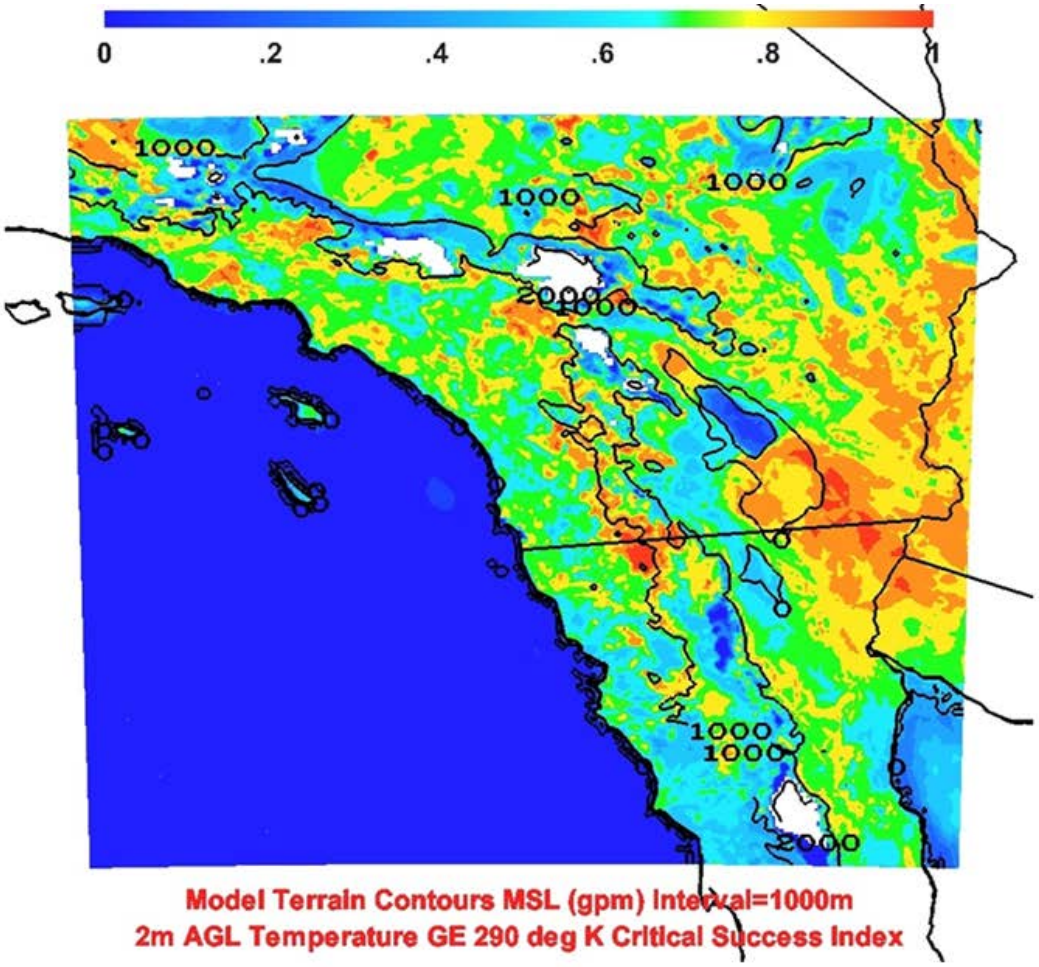

Fig. 2 CSI for 2-m AGL TMP at GE 290 K

Considering the factors that affect the CSI, the occurrence of near-zero CSI over the ocean may be related to the lack of a significant number of occurrences of TMP GE 290 K, since this TMP is at the highest part of the range for the entire domain and the likelihood for this to occur over the ocean is lower. More analysis of additional scores and statistics is needed to discern whether this is the case or there is another reason.

A display of the spatial distribution of the CSI for RH is shown in Fig. 3.

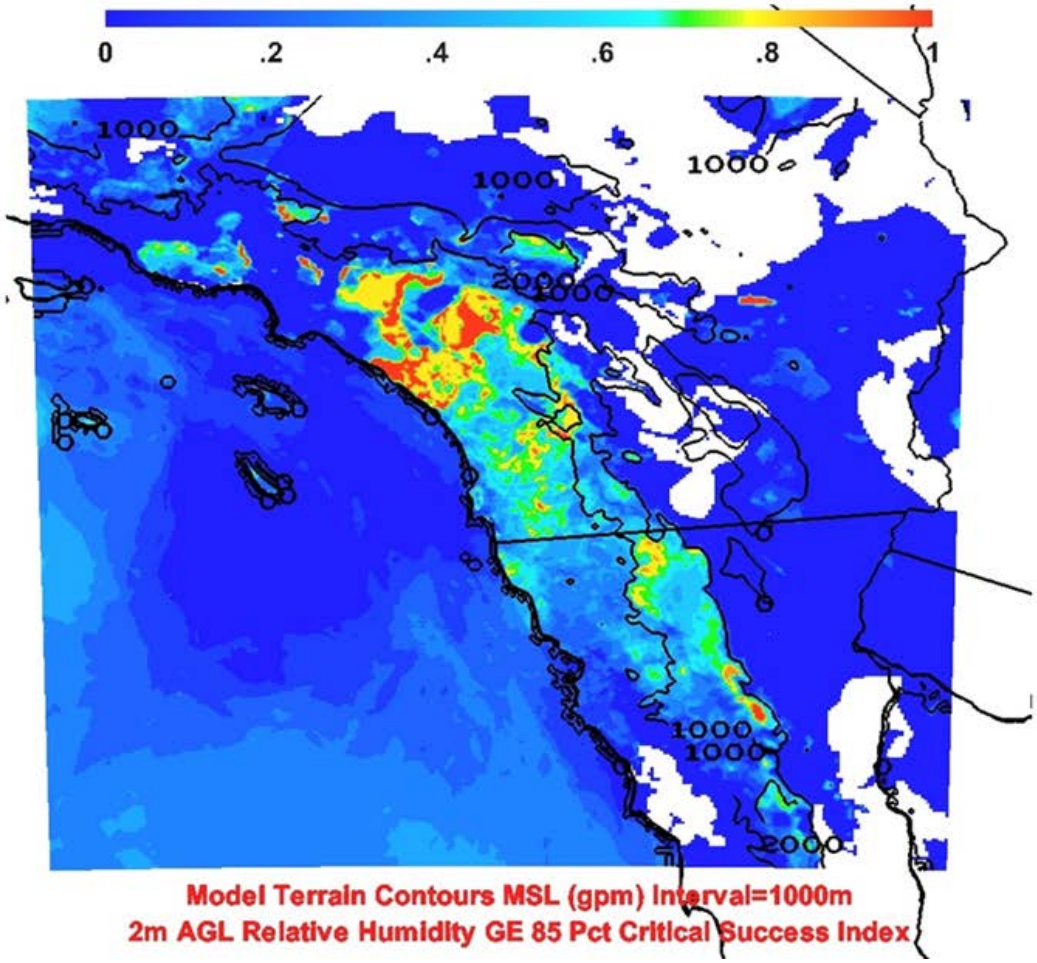

Fig. 3 CSI for 2-m AGL RH at GE 85%

The CSI for RH shows scoring over a large portion of the domain, but there is a significant part of the domain over interior land areas that has no scoring—as indicated by the white color—due to no occurrences of RH at GE 85%. There is a limited area of moderate to high CSI located inland in the coastal zone and lower-elevation areas, while a significant portion of the land area has low CSI. The CSI over the ocean is low to moderate.

The distribution of low to moderate CSI lies roughly equally between land areas and ocean areas. Over the water and some coastal areas (and perhaps some of the higher-elevation areas) the expectation of occurrences of RH at GE 85% is higher; thus, insufficient numbers of occurrences are not considered the cause of low scores in these areas. Analysis of additional scores and statistics is needed to understand the reasons for the poor performance in these areas.

A display of the spatial distribution of the CSI for WIND is shown in Fig. 4.

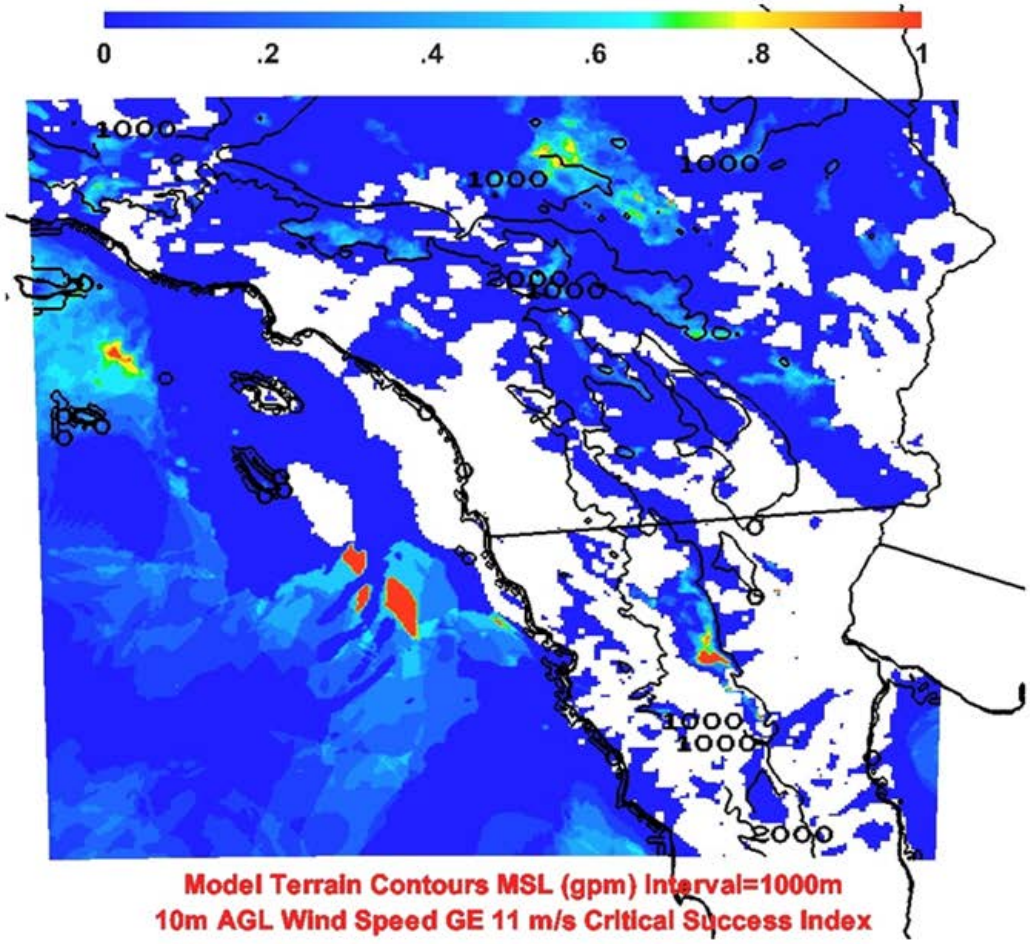

Fig. 4 CSI for 10-m AGL WIND at GE 11 m/s

The CSI for WIND shows that low to moderate CSI dominates the domain with higher CSI in isolated, small areas over the ocean. There are also extensive areas with no event occurrences for scoring, as indicated by the white color.

The poor performance in most of the domain may be related to low numbers of occurrences of WIND at GE 11 m/s. Analysis of additional scores and statistics is needed to understand the reasons for the overall poor performance.

4.1.2 Apply FBIAS for Assessment of the WRE–N

A display of the spatial distribution of the FBIAS for TMP is shown in Fig. 5. The FBIAS for TMP shows scoring over the entire domain with the exception of high-elevation locations in the mountains in the central portion of the domain (as indicated by white coloring), which represents no event occurrences for calculating scores. Most of the land areas have little bias with a notable exception in higher terrain in Mexico, where there are areas with a tendency to overforecast TMPs GE 290 K. The Salton Sea appears to be an area of anomalous underforecasting of the event for the reasons previously mentioned. The oceanic areas are homogeneous and also have a tendency for underforecasting of the event.

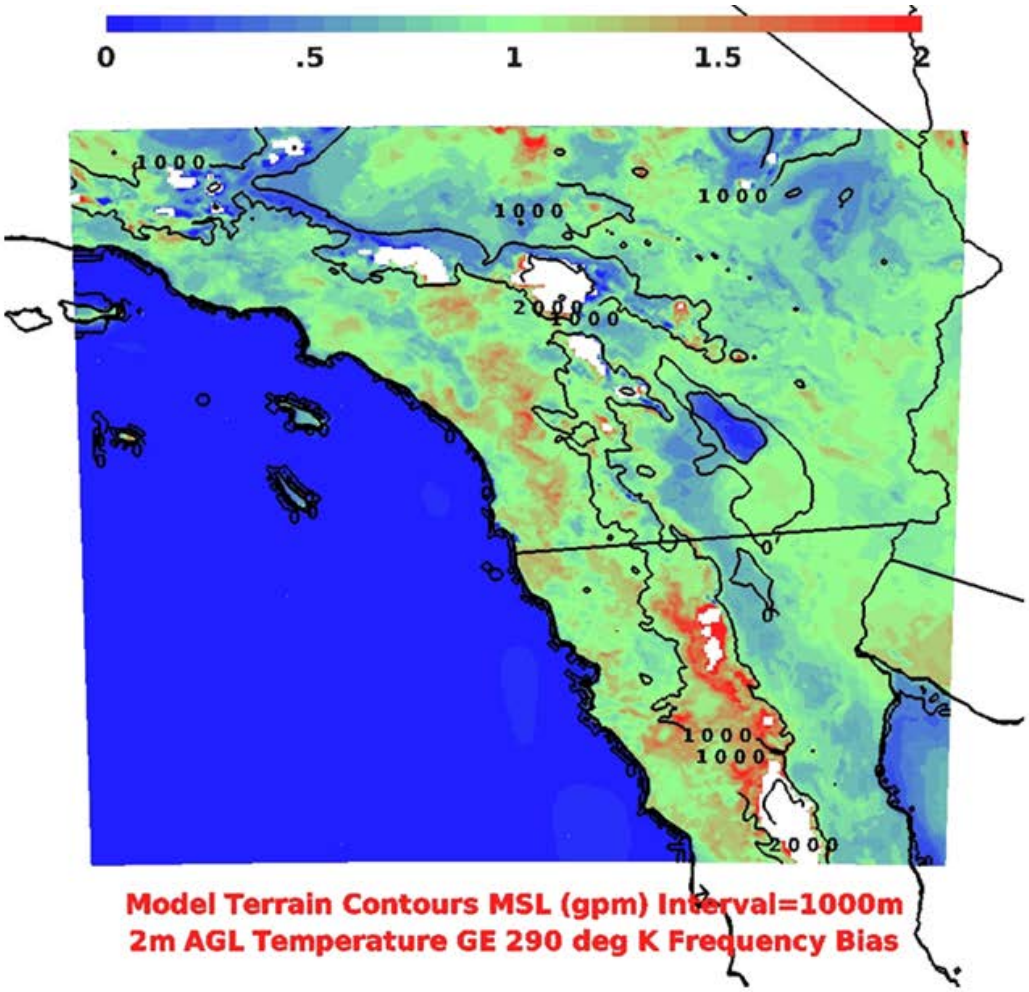

Fig. 5 Frequency bias for 2-m AGL TMP at GE 290 K

Noting a uniform pattern of underforecasting over the oceans coinciding with a similar uniform pattern of low CSI, again, this may be related to the lack of occurrences of TMP GE 290 K: this TMP is at the highest part of the range for the

entire domain and the likelihood for this to occur over the ocean is low. The extensive areas of small bias over land show good forecast skill, but not all of this area showed equally good CSI scores. The lower-elevation areas in the eastern part of Southern California are where higher CSI coincided with little or no bias and reflect good performance by the model. More analysis of additional scores and statistics is needed to better understand the uniform pattern of the underforecasting tendency over the ocean.

A display of the spatial distribution of the FBIAS for RH is shown in Fig. 6.

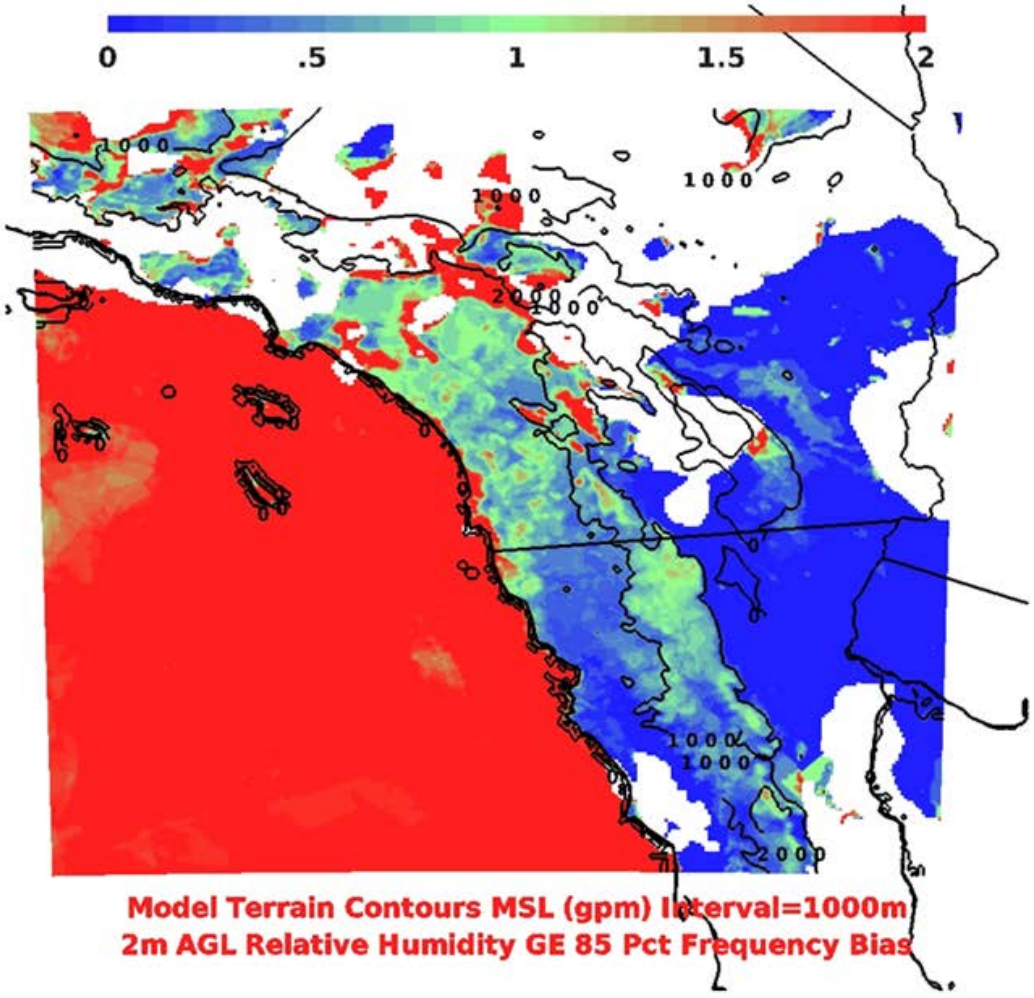

Fig. 6 Frequency bias for 2-m AGL RH at GE 85%

The FBIAS for RH at GE 85% gives a limited assessment of the model over land. There are significant areas over land with no apparent scoring, as indicated by the white color. Where there is scoring over land, there are areas of underforecasting bias and little to no bias over coastal areas and lower elevations of the

southeastern part of the domain. Over the ocean, the prevailing tendency is for overforecasting of RH.

The areas with no FBIAS scoring are most likely the result of the lack of occurrences of RH at GE 85%. Analysis of additional scores and statistics is needed to understand the reasons for the significant tendency for overforecasting of RH over the ocean.

A display of the spatial distribution of the FBIAS for 10-m AGL WIND is shown in Fig. 7.

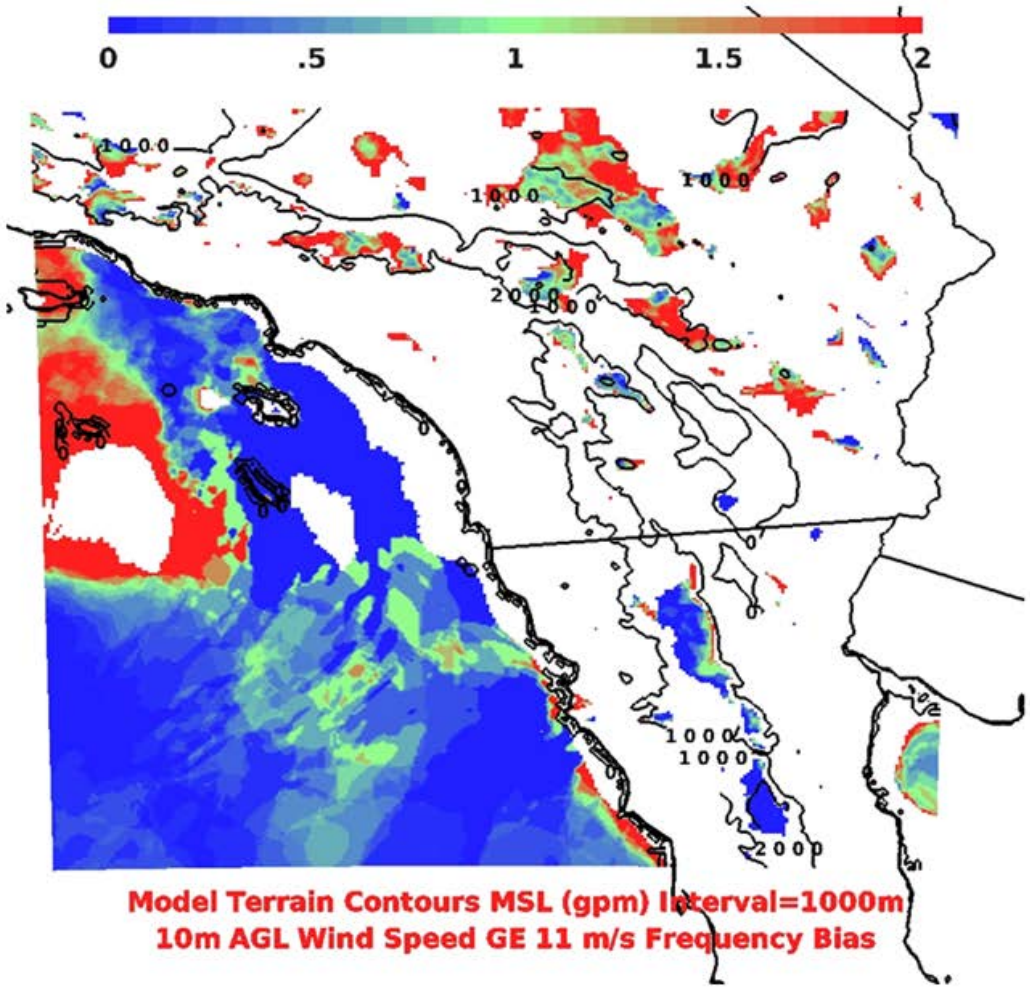

Fig. 7 Frequency bias for 10-m AGL WIND at GE 11 m/s

The FBIAS for WIND at GE 11 m/s shows a very limited assessment of the model. There are extensive areas with no scoring over land. The limited areas where there is scoring show small bias over higher-elevations locations in the United States and a notable underforecast tendency at higher elevations in Mexico. Over the ocean, there is better coverage of scoring with large areas of

underforecast bias surrounding a significant area of overforecast bias and with limited areas of little or no bias.

The areas with no FBIAS scoring are most likely the result of the lack of occurrences of WIND at GE 11 m/s. Analysis of additional scores and statistics is needed to understand the reasons for the limited amount of scoring.

4.1.3 Compare CSI and FBIAS Results for WRE–N with, without FDDA

MET Series-Analysis was run using output from the WRE–N that was produced without the FDDA to provide a basis for comparison of categorical skill scores with the WRE–N that was produced with the FDDA. Figures 8 and 9 display the spatial distribution of the CSI and FBIAS for TMP at GE 290 K and WIND at GE 11 m/s for both runs of the WRE–N.

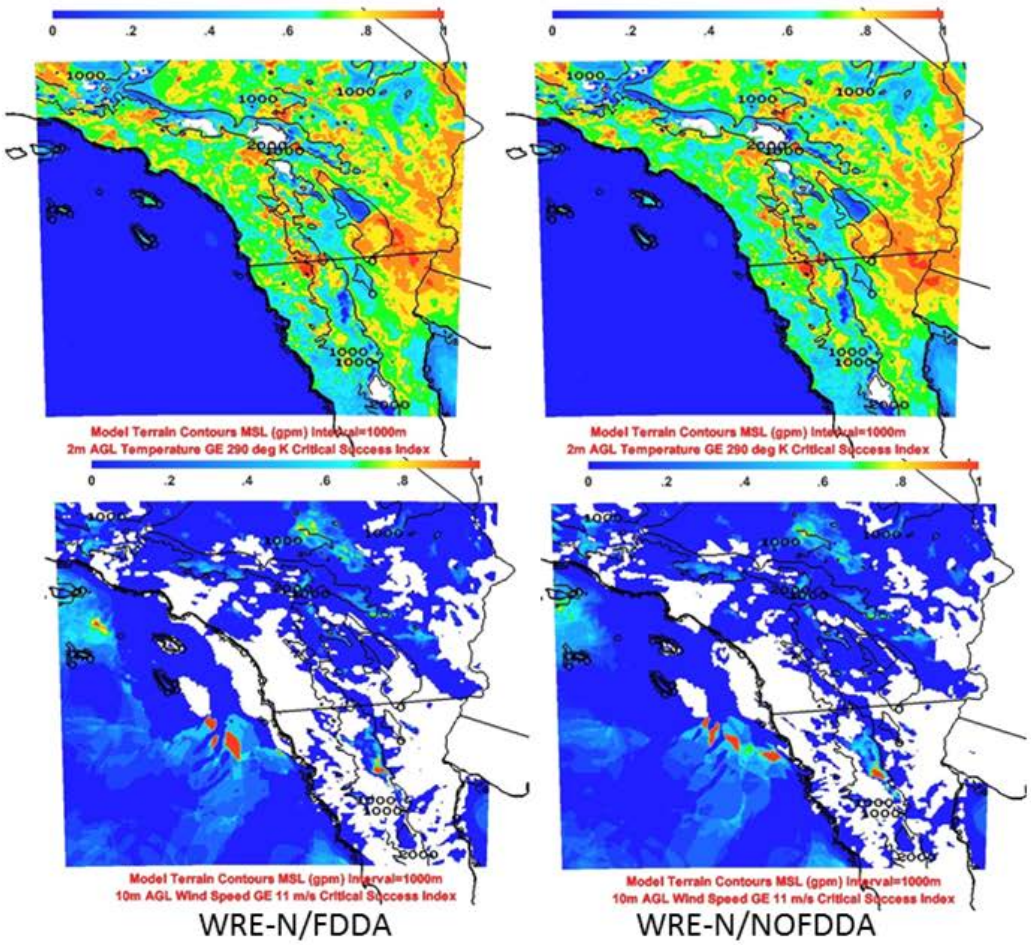

Fig. 8 CSI for 2-m AGL TMP at GE 290 K and WIND at GE 11 m/s for WRE–N with and without FDDA

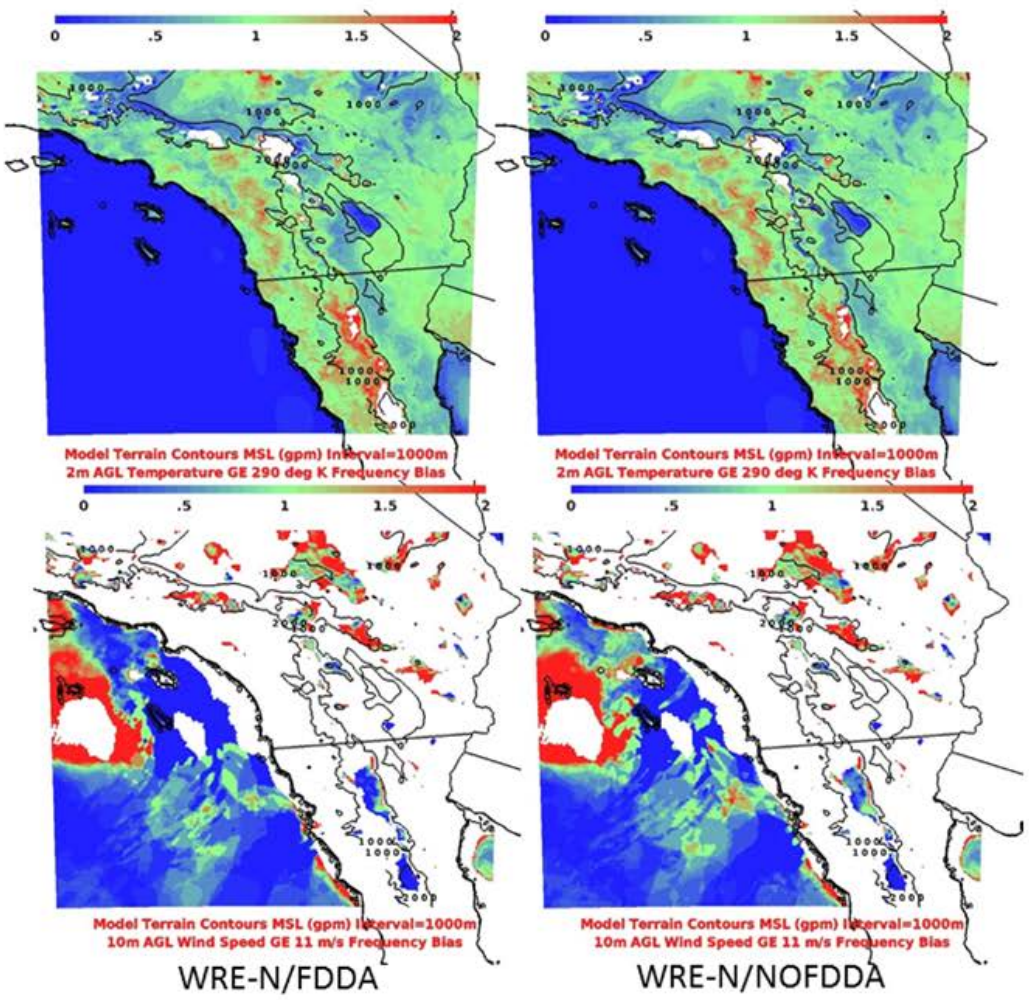

Fig. 9 Frequency bias for 2-m AGL TMP at GE 290 K and WIND at GE 11 m/s for WRE–N with and without FDDA

The purpose of running the WRE–N with FDDA is to improve the quality of the forecasts. The CSI and FBIAS were computed from forecasts generated by the WRE–N with and without FDDA to quantify the quality of each model run so a comparison could be made to determine the value added by running the model with FDDA. The CSI and FBIAS for both temperature and WIND show there is little apparent difference between the spatial distributions of the scores for both runs.

4.1.4 Summary of Application of Categorical Verification Techniques

The frequency of occurrence of forecast events determined by the application of a threshold to a continuous variable field varies over the domain and affects the CSI and FBIAS scores in a way that may give a misleading assessment of the model's ability to forecast objects. Analysis of more scores and contingency-table

statistics is needed to improve assessments of the ability of the model to forecast objects. Improved assessments of this aspect of model performance will lead to model improvements to enable better prediction of objects, which will, in turn, translate into better TDA weather-impact predictions.

The accuracy of the model judged from the scores varies considerably over the domain due to a combination of terrain characteristics and mesoscale variations in the air-mass characteristics. Analysis of more scores and contingency-table statistics is needed to better relate them to terrain and air-mass characteristics. The implications of this variability suggest that weather impacts on Army systems and missions vary considerably in space.

The value-added use of FDDA as judged from spatial displays of categorical scores and statistics is difficult to quantify. There are areas where the patterns of the scores computed with and without FDDA vary slightly over space, which implies the FDDA's value is a function of the terrain and/or mesoscale variations in air-mass characteristics present over the domain. Analysis of such differences using other approaches may provide more insight as to their causes.

The selection of the threshold to be used for generation of categorical verification scores and statistics will directly impact the extent of useful scores and statistics over the domain. Thus, it is important to include actual system and mission thresholds to more accurately assess the ability of the model to predict objects that are meaningful to the Army. That said, the use of actual thresholds will significantly reduce the number of locations and time periods in which the atmospheric conditions can provide the range of variable values that encompass these thresholds. The impact of these 2 situations—each at odds with the other—has to be judged with the understanding that meaningful conclusions about model performance can only come from the analysis of large numbers of cases. So, there is a tradeoff between analysis of data sets for fewer cases where tactically significant thresholds can be applied and data sets that were developed using thresholds defined by using the actual ranges of the variables present over the domain. The former presents challenges due to lack of statistically significant numbers of cases; the latter presents a challenge of limited application for assessment of the ability of models to forecast objects using mission- and system-specific thresholds.

4.2 Analysis of MET MODE Results

Traditional grid-versus-grid forecast-verification scores, such as CSI and FBIAS, provide a simple, straightforward picture of forecast quality—but, they offer very little diagnostic information, which is essential for modelers as well as model

users to better understand model performance. Feature- or object-based forecast-evaluation methods such as MODE were designed to fill the gap so that the reasons why a particular forecast is good or bad can be inferred. Traditionally, MODE has only been applied to sporadic, discontinuous fields such as precipitation, since it is natural to treat storms that produce precipitation as objects. There was no need to apply MODE to continuous variables such as TMP, RH, or WIND since continuous variables were normally verified against station observations using grid-to-point methods. However, a unique Army need to evaluate continuous variables as objects was identified when we consider that some Army TDAs, such as MyWIDA, employ thresholds on continuous variables such as TMP, RH, and WIND to identify potentially hazardous regions for Army operations. When such thresholds are applied by TDAs on a continuous field, an object is automatically created within that field. Therefore, it is desirable to know, for example, how well an object defined by WIND over 15 m/s in forecast is matched to its corresponding observed object. In other words, we strive to understand how the TDA's warning area is affected by forecast accuracy. Literature review as of this writing suggests our approach using MODE applied to continuous variables is unique. Lessons learned from this study will lay a solid foundation for future evaluations of the effectiveness of TDAs using meteorological data as their input.

The same method developed by Cai and Dumais (2015) for precipitation objects has been applied to surface TMP, RH, and WIND for the 5 case days described in Section 2 as a proof-of-concept study. Two statistics, the total area and total number of objects, were compiled as a function of forecast lead time. Since there was a 6-h preforecast period when FDDA was applied, the true free forecast starts at the 6-h lead time in the MODE analysis. The total area and total number of TMP, RH, and WIND objects with and without FDDA are shown in Figs. 10, 11, and 12, respectively. (Table 6 list the thresholds for TMP, RH, and WIND.)

Figure 10 shows an overforecast of total number of TMP objects (~100%) and an underforecast of total area of TMP objects (~25%), implying too many small objects in the forecast. Notably, the model did a great job at forecasting the trend of the total number and area of TMP objects, which corroborates what other researchers have determined (e.g., Wilson et al. 2010). Finally, Fig. 10 also seems to suggest that FDDA did not have noticeable impact on the results, which appears counter-intuitive and needs further investigation. One possible explanation: Because FDDA was performed on a point basis, its impact on the TMP objects, which are usually rather large in size and contain many grid points, is therefore limited unless the assimilated data points happened to be on the boundary of an object.

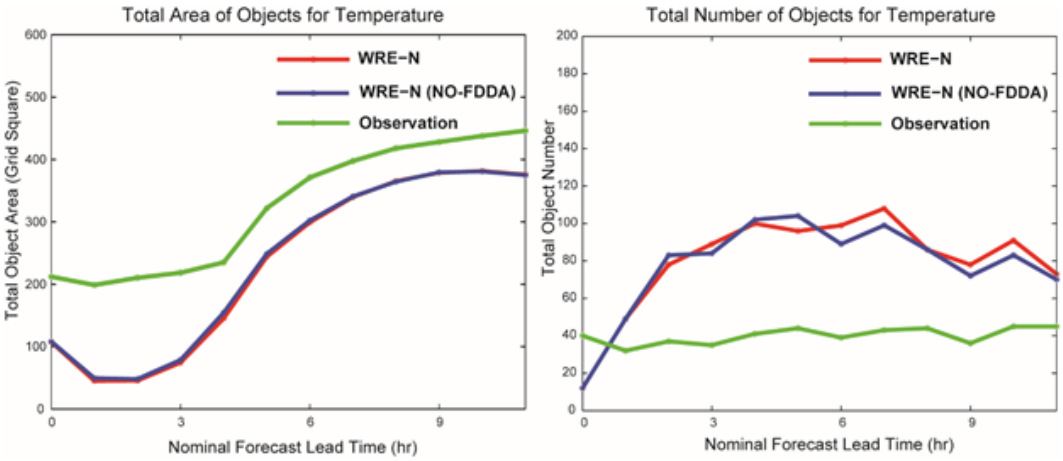

Fig. 10 Total area (left) and total number (right) of TMP objects compared to observations (green) for the WRE–N with FDDA (red) and without FDDA (blue) as a function of forecast lead time for the 5 case days described in Section 2.3. The TMP threshold used to identify objects in both forecast and observation is 290 K.

The total area and total number of RH objects compared to observations are shown in Fig. 11. Similar to TMP, the difference with and without FDDA is small, although the total number of objects without FDDA seems slightly better than with FDDA. Both the total number and total area of objects were overforecasted (approximately a factor of 2 for total number of objects and ~50% for total area of objects, respectively), which could have significant implications to TDAs.

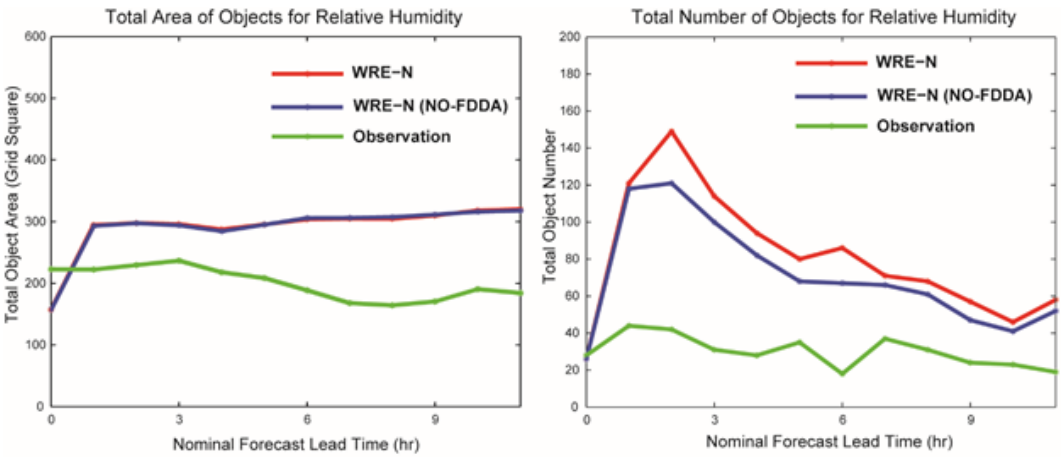

Fig. 11 Total area (left) and total number (right) of RH objects compared to observations (green) for the WRE–N with FDDA (red) and without FDDA (blue) as a function of forecast lead time for the 5 case days. The RH threshold used to identify objects in both forecast and observation is 85%.

Finally, the total area and total number of WIND objects are shown in Fig. 12. An overforecast of approximately a factor of 3 for the total number of objects was

noted in Fig. 12 while the total area of WIND objects was close to observations. This implies there are many small objects in the forecast compared to observations. Again, consistent with Figs. 10 and 11, the impact of FDDA seems insignificant, and the model did a good job at forecasting trend.

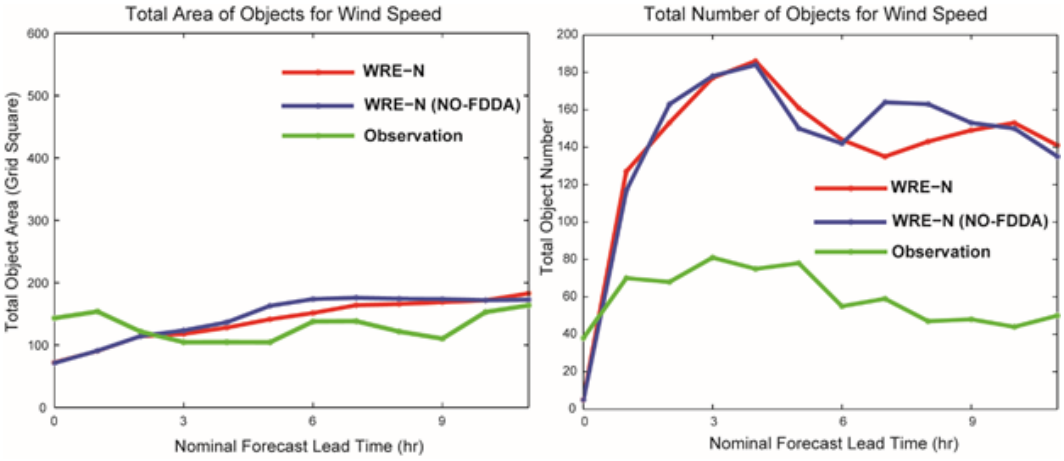

Fig. 12 Total area (left) and total number (right) of WIND objects compared to observations (green) for the WRE–N with FDDA (red) and without FDDA (blue) as a function of forecast lead time for the 5 case days. The WIND threshold used to identify objects in both forecast and observation is 11 m s^{-1}.

In summary, the model appears to have the lowest bias in terms of the total area of WIND objects, while it tends to underforecast the total area of TMP objects (~25%) but overforecast the total area of RH objects (~50%). Gaining a general idea of the bias of the model forecast could be beneficial for estimating the impact of forecast accuracy on TDAs used in Army operations.

This research serves as a proof of concept for using object-based forecast-evaluation tools such as MODE to assess the forecast that will be fed into a TDA. Thus, this preliminary study can be improved in many ways. For example, we should greatly expand the number of cases, trying a number of different thresholds—especially the thresholds that are meaningful for TDAs such as MyWIDA. Also, we could analyze more meteorological variables and compute more object attributes such as object-size distribution, just to name a few possibilities. The ultimate goal is to gauge the impact of forecast accuracy on TDAs used in Army operations, and we still have a long way to go. Hopefully this study will serve as a springboard to spearhead the efforts in that direction.

5. Conclusion and Final Comments

We have found the traditional method for verification of categorical forecasts offers a straightforward approach to assess the ability of the model to predict objects defined by the application of a threshold to a spatial forecast of a continuous variable. This study demonstrated 1) the applicability of the MET Series-Analysis tool for generating spatially distributed, categorical contingency-table statistics and scores for continuous meteorological variable fields and 2) that the CSI and FBIAS statistics will provide a limited assessment of model accuracy. However, due to the high spatial variability of these 2 statistics, analysis of additional scores and more cases is necessary. One reason for this is the choice of the value of the threshold: If the threshold is at the high end of the full range of the variable, there will be areas where no events will occur, which limits the area where scores can be calculated. Another reason for the limited assessment is the restriction of the analysis to only 2 statistics. There are numerous contingency-table scores and statistics that can be calculated and, when analyzed together, may reveal more information about model performance and provide a background to support more understanding of all the scores and statistics. We believe a more comprehensive approach of combining results from the traditional methods with those generated from the application of nontraditional object-based methods is best for an assessment of the skill of the model in predicting fields of a continuous variable that have been filtered by a threshold. Judging from the complexity of the spatial distribution of the CSI and the FBIAS, this more rigorous approach will certainly require a large amount of data so that statistically significant results can be obtained. In addition, it is found that the quality of the gridded observation data sets has an impact on the quality of the scores and statistics generated using the categorical method.

This preliminary study also documented the first attempt of applying an object-based forecast-evaluation method (i.e., MODE) to continuous meteorological variables. To the best of our knowledge, this novel approach has never been done before. Considering the Army TDAs mostly rely upon critical thresholds in continuous variables such as temperature, relative humidity, and wind speed to issue warnings that might affect Army operations, it is imperative to evaluate the impact of forecast accuracy on Army TDAs. By employing both traditional and nontraditional forecast-evaluation methods (such as those demonstrated in this study), a more complete picture of model-forecast performance can be gleaned by analyzing large amount of forecast data. Heading into that direction, future work will focus on more statistics and, most importantly, more cases so that statistically significant results can be obtained.

Finally, a Geographic Information System, which for the atmospheric sciences has not been extensively used, should be exploited for its ability to contextualize and analyze geospatial information such as 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 considerable promise of becoming an important new tool that, in addition to the methods described in this study, 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.

Cai H, Dumais RE. Object-based evaluation of a numerical weather prediction model's performance through forecast storm characteristic analysis. Wea Forec. 2015;30:1451–1468.

Casati B, Wilson LJ, Stephenson DB, Nurmi P, Ghelli A, Pocernich M, Damrath U, Ebert EE, Brown BG, Mason S. Forecast verification: current status and future directions. Meteo App. 2008;15(1):3–18.

Chen F, Dudhia J. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part II: preliminary model validation. Mon Wea Rev. 2001a;129:587–604.

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 Wea Rev. 2001b;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]. http://ams.confex.com/ams/pdfpapers/104773.pdf.

De Pondeca MSFV, Manikin GS, DiMego G, Benjamin SG, Parrish DF, Purser RJ, Wu W-S, Horel JD, Myrick DT, Lin Y, et al. 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, 2009 Jun 23–26; Boulder (CO).

Developmental Testbed Center. MET online tutorial for METv3.0: COPYGB functionality. Boulder (CO): National Oceanic and Atmospheric Administration; [accessed 2016 Jul 27]. http://www.dtcenter.org/met/users/support/online_tutorial/METv3.0/copygb/index.php.

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 observation 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 June. Report No.: Report ARL-TN-0546.

Dumais RE, Raby JW, Wang Y, Raby YR, Knapp D. Performance assessment of the three-dimensional wind field Weather Running Estimate-Nowcast and the three-dimensional wind field Air Force Weather Agency weather research and forecasting wind forecasts. White Sands Missile Range (NM): Army Research Laboratory (US); 2012 Dec. Report No.: ARL-TN-0514.

Dumais Jr RE, Henmi T, Passner J, Jameson T, Haines P, Knapp D. A mesoscale modeling system developed for the US 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, et al. Progress and challenges in forecast verification. Meteo App. 2013;20(2):130–139.

[EMC] Environmental Modeling Center. The GFS atmospheric model. Washington (DC): National Weather Service–National Centers for Environmental Prediction; 2003 Nov. NCEP Office Note No.: 442.

Google Earth. Mountain View (CA); 2016 [accessed 2016 Aug 24]. http://maps.google.com/help/terms_maps.html.

Hong S-Y, 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. Forecast verification: a practitioner's guide in atmospheric science. 2nd ed. Hoboken (NJ): John Wiley and Sons; 2012.

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. Paper presented at: 6th WRF/15th MM5 Users' Workshop, National Center for Atmospheric Research; 2005 Jun 27–30; 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.

Murray D, McWhirter J, Wier S, Emmerson S. The integrated data viewer—a web-enabled application for scientific analysis and visualization. Paper presented at: 19th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology; 2003.

[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.

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 Meeting of the American Meteorological Society, 23rd Conference on Probability and Statistics in the Atmospheric Sciences; 2016 Jan; New Orleans (LA).

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; Burlington (VT).

Stauffer DR, Seaman NL. Multiscale four-dimensional data assimilation. J App Meteo. 1994;33:416–434.

Vaucher G, Raby J. Assessing high-resolution weather research and forecasting (WRF) forecasts using an object-based diagnostic evaluation. White Sands Missile Range (NM): Army Research Laboratory (US); 2014 Feb. Report No.: ARL-TR-6843.

Wilks DS. Statistical methods in the atmospheric sciences. 3rd ed. Oxford (England): Academic Press; 2011.

Wilson JW, Feng Y, Chen M, Roberts RD. Nowcasting challenges during the Beijing Olympics: successes, failures, and implications for future nowcasting systems. Wea Forec. 2010;25:1691–1714.

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
CSI	critical success index
Deg K	degrees Kelvin
DPT	dew-point temperature
FBIAS	frequency bias
FDDA	Four-Dimensional Data Assimilation
GE	greater than or equal to
GFS	Global Forecast System
GRIB	Gridded Binary format, edition 1
GRIB2	Gridded Binary format, edition 2
GSD	Global Systems Division
hPa	Hectopascal
HRRR	High Resolution Rapid Refresh
LAPS	Local Analysis and Prediction System
MADIS	Meteorological Assimilation Data Ingest System
MET	Model Evaluation Tools
MODE	Method for Object-Based Diagnostic Evaluation
MYJ	Mellor–Yamada–Janjic
MyWIDA	My Weather Impacts Decision Aid
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration

NWP	Numerical Weather Prediction
NetCDF	Network Common Data Form
PBL	planetary boundary layer
RH	relative humidity
RRTM	Rapid Radiative Transfer Model
RTMA	Real-Time Mesoscale Analysis
TAMDAR	Tropospheric Airborne Meteorological Data Reporting
TDA	Tactical Decision Aid
TKE	turbulent kinetic energy
TMP	temperature
UTC	Coordinated Universal Time
WIND	wind speed
WRE–N	Weather Running Estimate–Nowcast
WRF	Weather Research and Forecasting
WRF–ARW	Weather Research and Forecasting, Advanced Research WRF

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDF)	DIRECTOR US ARMY RESEARCH LAB RDRL CIO L IMAL HRA MAIL & RECORDS MGMT
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
11 (PDF)	US ARMY RSRCH LAB 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
1 (PDF)	US ARMY RSRCH LAB ATTN RDRL CIE P CLARK
3 (PDF)	US ARMY RSRCH LAB ATTN RDRL CIE D R RANDALL S O'BRIEN J JOHNSON
1 (PDF)	US NAVY RSRCH LAB DR J MCLAY
1 (PDF)	US AIR FORCE R CRAIG
1 (PDF)	DCGS-A WX SVCS J CARROLL
3 (PDF)	UCAR T FOWLER J H GOTWAY B BROWN
1 (PDF)	USAICOE J STALEY