

9-10-1995

An Improved Method For Determining and Characterizing Alignments of Point-Like Features and Its Implications For the Pinacate Volcanic Field, Sonora, Mexico

Tim Lutz

West Chester University, tlutz@wcupa.edu

J T. Gutmann

Follow this and additional works at: http://digitalcommons.wcupa.edu/geol_facpub



Part of the [Geology Commons](#)

Recommended Citation

Lutz, T., & Gutmann, J. T. (1995). An Improved Method For Determining and Characterizing Alignments of Point-Like Features and Its Implications For the Pinacate Volcanic Field, Sonora, Mexico. *Journal of Geophysical Research-Solid Earth*, 100, 17659-17670.
<http://dx.doi.org/10.1029/95JB01058>

This Article is brought to you for free and open access by the Geology & Astronomy at Digital Commons @ West Chester University. It has been accepted for inclusion in Geology & Astronomy Faculty Publications by an authorized administrator of Digital Commons @ West Chester University. For more information, please contact wccressler@wcupa.edu.

Assessing the Short-Term Forecast Capability of Nonstandardized Surface Observations Using the National Digital Forecast Database (NDFD)

JOBY L. HILLIKER

Department of Geology and Astronomy, West Chester University, West Chester, Pennsylvania

GIRISH AKASAPU

Department of Computer Science, West Chester University, West Chester, Pennsylvania

GEORGE S. YOUNG

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

(Manuscript received 7 October 2008, in final form 15 February 2010)

ABSTRACT

The number of surface observations from nonstandardized networks across the United States has appreciably increased the last several years. Automated Weather Services, Inc. (AWS), maintains one example of this type of network offering nonstandardized observations for ~8000 sites. The present study assesses the utility of such a network to improve short-term (i.e., lead times <12 h) National Digital Forecast Database (NDFD) forecasts for three parameters most relevant to the energy industry—temperature, dewpoint, and wind speed. A 1-yr sample of 13 AWS sites is chosen to evaluate the magnitude of forecast improvement (skill) and influence of physical location (siting) on such improvements. Hourly predictions are generated using generalized additive modeling (GAM)—a nonlinear statistical equation incorporating a predetermined set of the most significant AWS and NDFD predictors. Two references are used for comparison: (i) persistence climatology (PC) forecasts and (ii) NDFD forecasts calibrated to the AWS sites (CNDFD). The skill, measured via the percent improvement (reduction) in the mean absolute error (MAE), of forecasts generated by the study's technique (CNDFD+) is comparable (<5%) to PC for lead times of 1–3 h for dewpoint and wind speed. Skill relative to PC slowly increases with lead time, with temperature exhibiting the greatest relative-to-PC skill (~30% at 12 h). When compared to baseline CNDFD forecasts, the MAE of the generated CNDFD+ forecasts is reduced 65% for temperature and dewpoint at the 1-h lead time. An exponential drop in improvement occurs for longer lead times. Wind speed improvements are notably less, with little skill (<5%) demonstrated for forecasts beyond 4 h. Overall, CNDFD+ forecasts have the greatest accuracy relative to CNDFD and PC for the middle (3–7 h) lead times tested in the study. Variations in CNDFD+ skill exist with respect to AWS location. Tested stations located in complex terrain generally exhibit greater skill relative to CNDFD than the 13-station average for temperature (and, to a lesser degree, dewpoint). Relative to PC, however, the same subset of stations exhibits skill *below* the 13-station average. No conclusive relationship can be made between CNDFD+ skill and the sample stations located near water.

1. Introduction

Improved short-term weather forecasting has been an increased focus in meteorology, in part because of the myriad of applications for accurate short-term (1–12 h, as defined here) forecasts at the economic level. As examples

from the transportation industry, short-term forecasts of thunderstorms and low ceiling are beneficial to aviation (Fabbian et al. 2007; Hilliker et al. 2007; Ghirardelli and Glahn 2010), while developed forecasting techniques of other weather parameters [e.g., blowing snow, as demonstrated in Baggaley and Hanesiak (2005)] can be applied to the short term to assist users in the trucking sector.

It is well established that improved short-term predictions can be generated by statistical analyses of observations (e.g., Vislocky and Fritsch 1997; Leyton and Fritsch 2003), or observations optimally combined with

Corresponding author address: Joby L. Hilliker, Department of Geology and Astronomy, West Chester University, 223 Merion Science Building, West Chester, PA 19383.
E-mail: jhilliker@wcupa.edu

model output, as in the model output statistics (MOS) approach (Glahn and Lowry 1972). This approach has been more recently expanded to the gridded MOS (Dallavalle and Glahn 2005; Glahn et al. 2008, 2009) and Localized Aviation MOS Program (LAMP; Ghirardelli 2005) products. To this end, these studies have shown the importance and robustness of the National Weather Service's (NWS) Automated Surface Observing System (ASOS) network (Nadolski 1998). As the network of gray dots in Fig. 1 shows, observing stations are located at airports, with observations more sparse in areas of complex terrain and/or outside of metropolitan areas. Moreover, the ~50-km average spacing between ASOS stations is too coarse for use in detecting subgrid model errors and making precise corrections for timing errors.

To address these and other data density issues with the nation's observational surface network, comparable networks and mesonetworks to ASOS with similar capabilities have been installed. The (NRC) National Research Council (NRC; 2008, their appendix B) provides a detailed list of these individual networks, which include the Road Weather Information System (RWIS; Quixote Transportation Technologies 2010), Remote Automated Weather Stations (RAWS; Zachariassen et al. 2003), and the National Resource Conservation Service's Snowpack Telemetry (SNOTEL) network (Serreze et al. 1999). One of the more prolific and publicly accessible data providers is the Citizen Weather Observer Program (CWOP; Helms 2005), which includes observations often taken from less expensive, nonstandardized weather stations owned by private citizens. The success of the plethora of these individual networks has fostered umbrella data providers, such as the Meteorological Assimilation and Data Ingest System (MADIS; Miller et al. 2005) and MesoWest (Horel et al. 2002), which aggregate observations from individual mesonetworks.

AWS Convergence Technologies, Inc., has recently installed an additional network of automated, nonstandardized instruments (AWS; AWS 2010). The black dots in Fig. 1 indicate the locations of ~8000 AWS (more commonly referred to by its trade name, "WeatherBug") sites across the United States, many outside of metropolitan areas. Most AWS stations are located atop schools and other public buildings. The AWS instrument package provides measurements of atmospheric variables including temperature, dewpoint, precipitation, and a 2-min wind average (AWS 2007).

There are, however, disadvantages of this network. Unlike the unrestricted access to many of the networks listed in NRC (2008), AWS data are propriety. Second, supplemental weather parameters such as cloud cover, visibility, and precipitation type cannot be observed by the AWS stations. Another consideration is local siting

observational error. Most AWS stations are installed atop buildings and thus there is no standardized height at which weather parameters are measured. Moreover, some variables can be biased (e.g., wind speed) according to the number and locations of surrounding buildings and trees. Nevertheless, AWS (2007) provides the station installer siting standards to optimize data quality.

Another source of observational error, discussed in Daley (1993), Myrick and Horel (2006), and Bondarenko et al. (2007), is representativeness error. This is a consideration for both nonstandardized and standardized mesonetworks, particularly where observations are sparse, and/or whose sites are located in areas (e.g., complex terrain, next to water) where the observed weather is not representative of the larger scale. Several studies (e.g., Benjamin et al. 1999; Liu and Rabier 2002; Janjić and Cohn 2006; Myrick and Horel 2008) have examined this complex issue as it has important implications with respect to model data assimilation.

Despite such limitations, nonstandardized supplemental networks can supply critical observations that fall in the mesoscale gaps between ASOS stations. As examples, RWISs provide vital data on critical transportation parameters (e.g., surface roadway temperature); fuel moisture from RAWS data is used for diagnosing fire danger, while CWOP wind speed data are crucial during severe weather. Moreover, these networks have been used to diagnose mesoscale weather phenomena (Ludwig et al. 2004; Geerts 2008), assess their impact on surface analyses (Myrick and Horel 2008), and explore their data robustness (Illton et al. 2008). Several studies have also explored the mesonetworks' short-term prognostic utility, for example, with respect to convection via the Oklahoma Mesonetwork (Hilliker et al. 2007), and sensible weather predictions by applying MesoWest observations (Hart et al. 2004).

Yet, there has been limited effort to test the forecast utility of nonstandardized surface observing networks, such as the AWS. Thus, the present study's goal is to test the application of such a network to improve short-term forecasts in the context of energy forecasting. Each hour, utility companies predict "load" (i.e., electricity usage) by consumers (Bolzern et al. 1982; Robinson 1997; Teisberg et al. 2005). For example, higher temperatures increase load during the summer since consumers are more likely to turn on air conditioners. If these forecasts can be improved, utility companies would be able to better anticipate load (Valor et al. 2001). This, in turn, would result in greater system reliability and profit since underestimating load forces utilities to purchase electricity at a much higher market price. Overestimating load, however, causes excess electricity to go to waste since it cannot be stored. Severe underforecasts can lead to system failure on a local or regional scale. Three variables pertinent to load forecasting

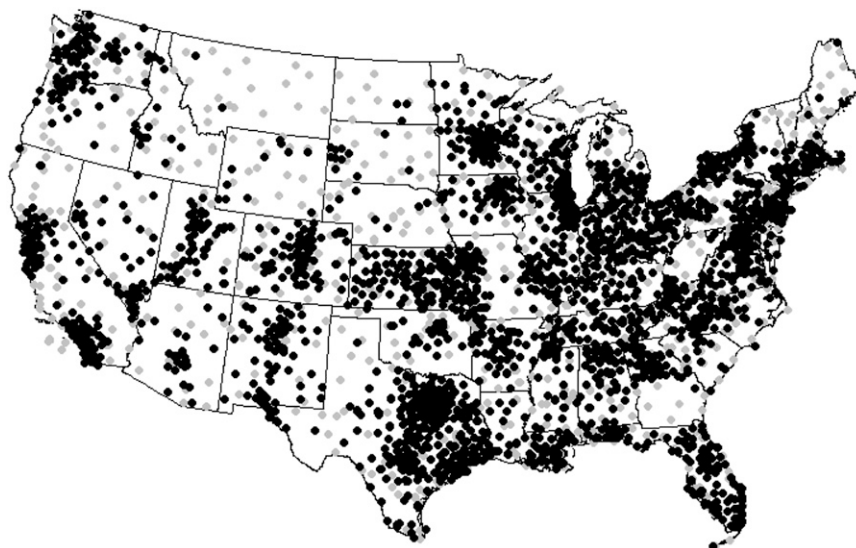


FIG. 1. Approximate surface observing sites composing the (a) ASOS network, as gray circles, and (b) AWS network, as black circles [modified from NWS (2006) and AWS (2010)].

are tested: temperature, dewpoint, and wind speed. Dewpoint is more relevant during the summer as it correlates with air-conditioning use, while wind speed is more relevant during the winter as it affects heating efficiency (Rothe et al. 2009; J. Quirk 2004, personal communication).

The methodology used to develop the forecasts is a regression approach, where the forecast variable (predictand) is determined using a statistical equation constructed by linking the predictand's most significant variables (predictors), determined ahead of time. Hart et al. (2004), Cosgrove and Sfanos (2004), Dallavalle et al. (2004), Hughes (2001), Schmeits et al. (2008), Hilliker et al. (1999), and Grover-Kopce and Fritsch (2003), among others, have shown success using this approach. The study differs from the latter two studies in that observations from surrounding stations are not considered. This work, instead, focuses on improving output from the National Digital Forecast Database (NDFD). Myrick and Horel (2006), Mollner (2005), and Dagostaro et al. (2004), as examples, have demonstrated the positive impact of the NDFD, particularly in short-term weather forecasting. In contrast, this work will demonstrate that observations from such a network can add value to the NDFD via calibrated NDFD 1–12-h predictions of relevant energy parameters. These calibrated NDFD (CNDFD) forecasts serve as a baseline, and allow for fair testing since CNDFD forecasts account for inherent differences between original NDFD forecasts and AWS instrumentation and location (e.g., those due to siting issues described previously). The forecasts generated by this study's technique will also be compared to persistence climatology (PC), a commonly used reference in short-term forecasting (Buell 1958;

Murphy 1992). Also known as conditional persistence, PC combines the well-known strength of applying persistence in short-term forecasting to knowledge of the variable's evolution for that particular time and station based on information contained in a historical dataset (Wilks 2006).

Finally, this study will also explore the magnitudes of improvement over CNDFD and PC forecasts as a function of AWS observing location. Forecasting over complex terrain is particularly challenging. Ruth et al. (2009) showed that gridded MOS scores were worse than station MOS scores over complex terrain; Myrick and Horel (2006) stated that larger temperature errors existed using Rapid Update Cycle (RUC) surface analyses and MesoWest data over the west than average temperature errors nationwide [from Benjamin et al. (2004), using RUC analyses and ASOS data]. Observations from the AWS network are hypothesized to provide the greatest improvement over CNDFD and PC forecasts for predictions at AWS sites located next to water or in areas of complex terrain, where subgrid terrain effects are expected to be large. This hypothesis is supported by Hart et al. (2004), which successfully applied MesoWest data to improve gridded model forecasts in the mountainous Utah terrain. Thus, results from the present study will support existing work by testing sites across different climates and terrains using a different network.

Section 2 presents the datasets used in this study and their quality control. The methodology and statistical design of the system are detailed in section 3. Section 4 is a summary of results from the dependent dataset, while section 5 presents results from the independent dataset. Section 6 summarizes the work's conclusions.

TABLE 1. Summary of AWS station identifiers tested with geographical data.

AWS identifier	Station name	Station location	Lat (°N)	Lon (°W)	Elev (m)
BRTTN	Bretton Woods Ski Resort	Bretton Woods, NH	44.26	71.44	499
LRAY1	Shenandoah National Park	Big Meadows, VA	38.52	78.44	1052
PHLRH	Rohm and Haas	Spring House, PA	40.18	75.20	98
RBBPH	Boardwalk Plaza Hotel	Rehoboth Beach, DE	38.72	75.08	16
WIVBT	WIVB-TV Station	Buffalo, NY	42.95	78.88	200
RCHSC	Seton Catholic High School	Richmond, IN	39.82	84.89	285
CHINM	P. Notebaert Nature Museum	Chicago, IL	41.92	87.63	179
BRRRD	Burr Ridge Middle School	Burr Ridge, IL	41.72	87.95	212
WNTRE	Wintergreen Mountain	Wintergreen, VA	37.92	78.94	1028
MTQCR	Crested Butte Mountain Resort	Crested Butte, CO	38.90	106.97	2843
MELBA	Dans Ferry Service	Melba, ID	43.35	116.60	853
SEASF	KING5 at SAFECO field	Seattle, WA	47.59	122.33	20
KPEAE	Peabody-Burns Elementary School	Peabody, KS	38.17	97.10	423

2. Datasets

Hourly AWS observations and NDFD forecasts were compiled for the 1-yr period from 15 April 2006 to 14 April 2007. Thirteen AWS stations, listed in Table 1, were tested. The test stations represent a variety of climates across the United States. For example, some station locations feature higher elevations [Crested Butte Mountain Resort (MTQCR) and Dans Ferry Service (MELBA), in the Rocky Mountains], proximity to a large body of water [KING5 at SAFECO field (SEASF), on the West Coast], or more homogeneous terrain with no obvious local effects (KPEAE, in Kansas).

NDFD forecasts are constructed and updated by the various National Weather Service Weather Forecast Offices (WFO). Output from operational computer models such as the RUC (Benjamin et al. 2004), North American Mesoscale (NAM; DiMego 2006), and short-range ensemble forecast (SREF; Du et al. 2009) system serves as the foundational field, and is mapped to a $2.5 \text{ km} \times 2.5 \text{ km}$ grid. Predictions are available at 3-h intervals (0000, 0300, 0600 UTC, etc.). The forecaster can modify the model field by applying other data, including climatic information, ASOS observations, and mesonet observations (Glahn and Ruth 2003). Ruth et al. (2009) have most recently stated that gridded MOS “should provide good guidance for preparing the NDFD.”

In addition to forecaster and WFO variations in NDFD generation, variations in human update frequency also exist. All WFOs update the NDFD grids at 0400 and 1600 LT—the two traditional major updates. Human adjustments to the forecast, particularly for the short term, between these cycles also occur. Compulsory 3-h NDFD updates at all WFOs, however, are anticipated in the future (D. Iovino 2009, personal communication). The final NDFD forecast is disseminated and merged with other WFO NDFD grids to form a national $5 \text{ km} \times 5 \text{ km}$ resolution product.

The NDFD data used in the study were obtained from the National Climatic Data Center’s (NCDC) online National Operational Model Archive and Distribution System (NOMADS). An accompanying decoder “probed” the data’s Gridded Binary (GRIB) format to retrieve NDFD forecasts, bilinearly interpolated to each AWS site’s latitude and longitude. In addition, although data files are available hourly, valid forecasts are available every 3 h (0000, 0300, 0600 UTC, etc.), consistent with WFO construction. The following two hours’ forecasts (e.g., those made at 0100 and 0200 UTC) are generally a repeat of the original forecast (the 0000 UTC forecast, to continue this example) until the next 3-hourly data point is available (0300 UTC, to complete the example).

Next, AWS and NDFD data that were missing or failed quality control (QC) checks were removed from analysis. This critical step ensures that the forecast system demonstrated in this study can extract the strongest statistical signals contained in the historic datasets. AWS employs an in-house QC technique for its data (J. Dutton 2007, personal communication); however, to verify the robustness of both datasets, tolerance limits from the NWS’s Quality Control and Monitoring Systems (QCMS) were applied (NWS 1993). Moreover, spatial consistency between AWS and collocated NDFD data was examined to reveal any obvious AWS siting issues. One flagged dataset was SEASF wind speed, which was consistently and anomalously lower ($\sim 0.6 \text{ m s}^{-1}$ average over the 1-yr sample) than surrounding ASOS observations and NDFD forecasts ($\sim 3.0 \text{ m s}^{-1}$ average over the same period). Thus, SEASF wind speed forecasts were not analyzed in the study. Nevertheless, the AWS and NDFD archives were generally robust, with $\sim 7\%$ of the data deemed bad or missing.

3. Statistical design of forecast system

Hourly deterministic forecasts of temperature, dew-point, and wind speed were made for each AWS station

for all forecast hours (0000–2300 UTC) for four lead times, ranging from 1 to 12 h. Table 2 summarizes the permutations of tested forecast hours, lead times, and resultant valid times. For example, a forecast made at 1900 UTC had four discrete lead times of 2, 5, 8, and 11 h, corresponding to valid times of 2100, 0000, 0300, and 0600 UTC, respectively. These particular valid times allow the most recently updated NDFD to be tested.

a. Generation of forecasts

Forecasts were generated by, first, linking the most significant predictors from the AWS and NDFD datasets using a generalized additive model (GAM) to form forecast equations. The equation's predictand, or forecast variable, is simply the AWS observation at the valid time. Generalized additive models are a nonparametric modeling technique that extends traditional multiple linear regression by objectively estimating the functional (curvative) relationship between the predictand and predictors (S-PLUS 2001; Vislocky and Fritsch 1995). As Vislocky and Fritsch (1995) summarized, "GAM fits a model as a sum Y of unspecified functions of the individual predictors X_n ,

$$Y \approx f_1(X_1) + f_2(X_2) + \cdots + f_n(X_n), \quad (1)$$

where the nonparametric functions f_i (for $i = 1, \dots, n$) are estimated from the data using smoothing operations (e.g., kernels, running means, splines). In this study, the cubic smoothing spline is employed. Vislocky and Fritsch (1995) showed that by using the cubic smoothing spline function in GAM, the mean square errors of short-term cloud cover, ceiling, and visibility forecasts were 3%–4% lower when compared to forecasts generated using multiple linear regression.

Figure 2 shows an example cubic smoothing spline from the study's dependent dataset. One of the predictands (Y)—1200 UTC wind speed at MELBA—was plotted against the predictor variable (X_1)—a 9-h NDFD forecast of MELBA wind speed. Note that the cubic spline captures the modest nonlinearity in the variables' relationship. For a detailed discussion on GAMs and the cubic smoothing spline, the reader is encouraged to review De Boor (1978), Hastie and Tibshirani (1990), and Vislocky and Fritsch (1995).

For each weather parameter, three equations were generated for each combination of forecast hour (of which there were 24, ranging from 0000 to 2300 UTC) and lead time (of which there were 4, ranging from 1 to 12 h, as explained above). The first member of each trio was that of the baseline-calibrated NDFD (CNDFD) forecasts. These one-predictor equations incorporated the original

TABLE 2. Exhaustive permutations of forecast hours (UTC without lead or trailing 00), lead times (h), and corresponding valid times (UTC without lead or trailing 00) tested. As an example, one 1900 UTC forecast has an 8-h lead time, or corresponding 0300 UTC valid time.

Forecast hour	Lead times (h)	Valid times, respectively
00	3, 6, 9, 12	03, 06, 09, 12
01	2, 5, 8, 11	03, 06, 09, 12
02	1, 4, 7, 10	03, 06, 09, 12
03	3, 6, 9, 12	06, 09, 12, 15
04	2, 5, 8, 11	06, 09, 12, 15
05	1, 4, 7, 10	06, 09, 12, 15
06	3, 6, 9, 12	09, 12, 15, 18
07	2, 5, 8, 11	09, 12, 15, 18
08	1, 4, 7, 10	09, 12, 15, 18
09	3, 6, 9, 12	12, 15, 18, 21
10	2, 5, 8, 11	12, 15, 18, 21
11	1, 4, 7, 10	12, 15, 18, 21
12	3, 6, 9, 12	15, 18, 21, 00
13	2, 5, 8, 11	15, 18, 21, 00
14	1, 4, 7, 10	15, 18, 21, 00
15	3, 6, 9, 12	18, 21, 00, 03
16	2, 5, 8, 11	18, 21, 00, 03
17	1, 4, 7, 10	18, 21, 00, 03
18	3, 6, 9, 12	21, 00, 03, 06
19	2, 5, 8, 11	21, 00, 03, 06
20	1, 4, 7, 10	21, 00, 03, 06
21	3, 6, 9, 12	00, 03, 06, 09
22	2, 5, 8, 11	00, 03, 06, 09
23	1, 4, 7, 10	00, 03, 06, 09

NDFD prediction corresponding to the forecast lead time. For example, the 1900 UTC CNDFD forecast equation for an 8-h lead time (i.e., 0300 UTC valid time) was

AWS 0300 UTC observation

$$\approx f(\text{NDFD forecast valid at 0300 UTC}), \quad (2)$$

where $f(\cdot)$ is the GAM nonparametric smoothing function discussed previously. The second member of each equation trio incorporated all predictors chosen by a stepwise regression method for that forecast hour. These equations were used to generate the final forecasts (referred to as CNDFD+, hereinafter). The 1900 UTC CNDFD+ forecast equation for an 8-h lead time was

AWS 0300 UTC observation

$$\approx f_1(X_1) + f_2(X_2) + \cdots + f_n(X_n), \quad (3)$$

where X_i (for $i = 1, \dots, n$) were the set of predictors chosen by stepwise regression (including both NDFD and AWS predictors), and f_i (for $i = 1, \dots, n$) as previously defined.

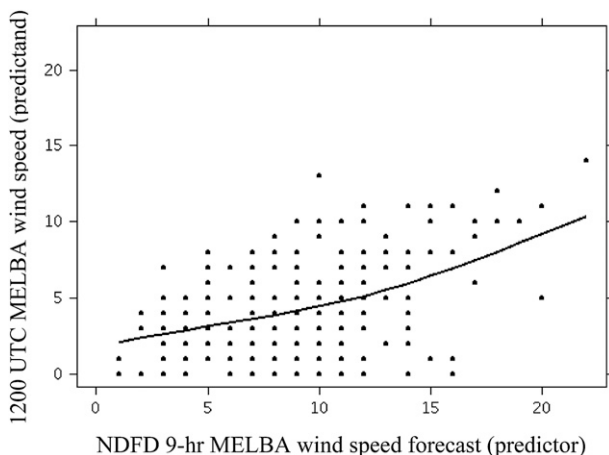


FIG. 2. Example cubic smoothing spline fit of observed 1200 UTC MELBA wind speed (y axis) vs its 9-h NDFD forecast (x axis).

The final equation in the trio constructed for each forecast hour, lead time, and predictand is a second reference set of forecasts to which those of the CNDFD+ can be compared—persistence climatology. The 1900 UTC PC forecast equation for an 8-h lead time (i.e., 0300 UTC valid time) is defined as

$$\begin{aligned} &\text{AWS 0300 UTC observation} \\ &\approx f(\text{AWS 1900 UTC observation}), \end{aligned} \quad (4)$$

where $f(\cdot)$ is previously defined.

b. Obtaining significant predictors

Table 3 lists the candidate predictors considered for short-term forecasting of temperature. Intuitively, the most promising variables to consider include the most recent (time = 0) temperature observation (T_0), as well as the most recent relative humidity, wind components, and past temperature observations. Equally as critical is considering the suite of NDFD 1–12-h forecasts. To continue the example above, the 1900 UTC forecast with an 8-h lead time considers, among other predictors, the NDFD forecast valid at 0300 UTC (annotated as NDFD_{7–9}, where “7–9” represents the block of lead times in which the forecast projection falls).

The AWS and NDFD data archives were divided as follows: days 1–23 of each month were dedicated to the larger dependent (or developmental) dataset from which the strongest predictors for short-term temperature, dewpoint, and wind speed forecasting were derived. The balance of the month was committed to the smaller independent dataset used to test the resultant forecast system.

TABLE 3. Description of candidate predictors tested and their notations.

Predictor	Notation
Current temperature	T_0
Temperature 1 h ago	T_{-1}
Temperature 2 h ago	T_{-2}
Current RH	RH_0
Current dewpoint	TD_0
Current wind direction	DD_0
Current wind speed	FF_0
Current precipitation (yes/no binary predictor)	P_0
Current U component of wind	U_0
Current V component of wind	V_0
NDFD 1–3-h forecast	$NDFD_{1-3}$
NDFD 4–6-h forecast	$NDFD_{4-6}$
NDFD 7–9-h forecast	$NDFD_{7-9}$
NDFD 10–12-h forecast	$NDFD_{10-12}$

The statistical software package S-PLUS was used to ascertain the set of optimal predictors, their f values (degree of linear association with the predictand), and ranking order (S-PLUS 2001). To obtain the optimal set of predictors, the “Efroymson” method was selected as the stepwise regression procedure (Efroymson 1960; Martin et al. 1963). This method is similar to a forward selection procedure in that a predictor is chosen based on its ability to produce independently the largest reduction in the residual sum of squares. However, when a new predictor is added to the subset, the Efroymson method determines if any of the previously selected predictors in the subset no longer contributes significantly to the modeled fit. If this is the case, the irrelevant predictor is eliminated from the regression equation.

In addition, the number of predictors to include in the forecast system depends on a prescribed “cutoff” f value (f_c). As specified in comparable studies, an f_c of 10 was applied here (Hilliker et al. 2007). Once the absolute value of the f value of the next significant predictor falls below f_c , no additional predictors are included, and the equation is finalized. Although it may be tempting to include many predictors to achieve the best modeled fit, the risk of “overfitting” increases. Overfitting is defined as only including predictors meaningful to the dependent dataset, which results in degraded performance when applied to the independent dataset. Additional information on stepwise regression, choosing f values, and overfitting can be found in Wilks (2006) and Neter et al. (1996).

4. Dependent data results

It is instructive to first explore which predictors from the dependant dataset were chosen by stepwise regression.

TABLE 4. A sampling of the strongest predictors as a function of forecast hour (UTC without lead or trailing 00) and lead time for temperature, dewpoint, and wind speed forecasting for CHINM. Predictors are listed in order of decreasing power. Predictor notation is referenced in Table 3.

Temperature				
Forecast hour	3-h lead time	6-h lead time	9-h lead time	12-h lead time
00	T_0 , NDFD ₇₋₉ T_{-1}	NDFD ₄₋₆ , T_0 U_0 , T_{-1}	NDFD ₄₋₆ , T_0 NDFD ₁₋₃ , U_0	NDFD ₁₀₋₁₂ , T_0 U_0
12	T_0 , NDFD ₄₋₆ T_{-1}	NDFD ₄₋₆ , T_0 , V_0	NDFD ₇₋₉ , V_0	NDFD ₇₋₉ , V_0
Dewpoint				
Forecast hour	3-h lead time	6-h lead time	9-h lead time	12-h lead time
00	TD ₀ , NDFD ₄₋₆	TD ₀ , NDFD ₄₋₆	NDFD ₇₋₉ , RH ₀	NDFD ₁₀₋₁₂ , RH ₀
12	TD ₀ , NDFD ₄₋₆	NDFD ₄₋₆ , TD ₀ U_0	NDFD ₇₋₉ , U_0 TD ₀	NDFD ₁₀₋₁₂ , U_0 T_0
Wind speed				
Forecast hour	3-h lead time	6-h lead time	9-h lead time	12-h lead time
00	NDFD ₁₋₃	NDFD ₄₋₆ , V_0	NDFD ₇₋₉	NDFD ₁₀₋₁₂ , RH ₀
12	NDFD ₁₋₃ , FF ₀	NDFD ₁₀₋₁₂	NDFD ₁₀₋₁₂	NDFD ₁₀₋₁₂

Table 4 shows a sampling of the final temperature, dewpoint, and wind speed predictors included in the forecast system for the P. Notebaert Nature Museum (CHINM). Predictors are listed in order of benefit, with the first predictor being the most highly correlated to the predictand. The nature, order, and number of predictors (typically, 2–4) were consistent for other AWS stations.

One of the most noteworthy results in the final predictor list is the tendency for the most recent AWS observation to be weighted more heavily for shorter lead times and NDFD output for longer lead times. In addition, the most beneficial NDFD predictors are typically those that correspond to the lead time (i.e., NDFD₇₋₉ for a 9-h lead time with dewpoint). There are exceptions (e.g., NDFD₄₋₆ for a 3-h 1200 UTC temperature forecast), however, that incorporate supplemental lead times. The additional independent information implied by the presence of multiple NDFD predictors may be the result of the system recognizing the benefit of the currently quasi-irregular human NDFD updates.

The limitations of short-term wind speed forecasting are also evident in Table 4. Even at the 3-h lead time, the most recent wind speed observation (FF₀) is generally absent. In fact, there is an overall lack of supplemental predictors for wind speed regardless of forecast or lead time. One additional result of note is the presence of the wind components (U_0 and V_0) as valuable predictors for forecasting temperature, and to a lesser degree dewpoint, for CHINM, a site located near Lake Michigan. A sample of the chosen predictors for MTQCR (see Table 5), the AWS site highest in elevation (~2800 m) from the sample, supports the importance of using

observations to modify NDFD forecasts, particularly where local effects dominate. The table is populated with more observational predictors, even through lead times of 12 h.

5. Independent data results

To assess CNDFD+ forecast quality, the mean absolute error (MAE) was compared for the CNDFD+, CNDFD, and PC forecasts¹ using the independent dataset, which yielded ~700 cases for each lead time. Supplemental metrics of forecast assessment could also be applied, as detailed in Wilks (2006). Figure 3 shows the MAE of CNDFD+, CNDFD, and PC temperature, dewpoint, and wind speed predictions as a function of lead time for a sampling of AWS sites. Results for each lead time were averaged over the eight (0000, 0300, 0600 UTC, etc.) valid times shown in Table 2.

The average CNDFD+ temperature MAE for CHINM, a typical example, was 0.6°C for a 1-h lead time and increased asymptotically to 1.8°C by 12 h. For comparison, baseline CNDFD forecast error was largely independent of lead time (MAE ~ 1.8°C for the 1–12-h lead times). Notable variations in CNDFD MAE with respect to valid time, however, exist. For example, for 12-h forecasts valid at 2100 UTC, the MAE was ~1.4°C, but increased to ~2.2°C for 12-h forecasts valid at 0900 UTC. Myrick and Horel (2006) hypothesize causes of the diurnal variation in NDFD error.

¹ Any wind speed forecast that was $<0 \text{ m s}^{-1}$ was truncated to 0 m s^{-1} .

TABLE 5. As in Table 4, but for MTQCR.

Temperature				
Forecast hour	3-h lead time	6-h lead time	9-h lead time	12-h lead time
00	$T_0, RH_0, T_{-2}, NDFD_{10-12}$	$T_0, NDFD_{7-9}$ RH_0, T_{-2}	$T_0, NDFD_{7-9}, TD_0, T_{-2}$	T_0, TD_0 $NDFD_{10-12}, T_{-2}$
12	$NDFD_{7-9}, T_0, P_0, NDFD_{1-3}, U_0, TD_0$	$NDFD_{7-9}, T_{-1}, NDFD_{1-3}, FF_0$	$NDFD_{7-9}, NDFD_{1-3}, T_{-2}$	$NDFD_{7-9}, RH_0$
Dewpoint				
Forecast hour	3-h lead time	6-h lead time	9-h lead time	12-h lead time
00	TD_0	$TD_0, NDFD_{10-12}$	$TD_0, NDFD_{10-12}$	$TD_0, NDFD_{10-12}$
12	TD_0, T_0, FF_0	TD_0, T_0	TD_0, RH_0	TD_0, RH_0 $NDFD_{10-12}$
Wind speed				
Forecast hour	3-h lead time	6-h lead time	9-h lead time	12-h lead time
00	$FF_{-2}, NDFD_{4-6}, FF_0$	FF_0, V_0 $NDFD_{7-9}, U_0$	$NDFD_{1-3}, FF_{-2}$	$NDFD_{7-9}, T_0$
12	U_0, V_0 $T_0, NDFD_{1-3}$	$NDFD_{7-9}, U_0$ T_0	$NDFD_{10-12}$	$NDFD_{10-12}$

The MAE of the reference PC forecasts was similar in magnitude to that of the CNDFD+ forecasts for the 1-h lead time, but exhibited an error rate much faster with lead time than corresponding CNDFD+ forecasts. By 5 h, PC forecasts were less accurate than CNDFD forecasts. Dewpoint results exhibited magnitudes and patterns similar to those of temperature, although for SEASF, PC forecast accuracy did not degrade as quickly with lead time as compared to CNDFD+ (see Fig. 3). By the 10-h lead time, PC predictions became less accurate than the baseline CNDFD forecasts. For wind speed, the CNDFD forecast error at the Bretton Woods Ski Resort (BRTTN) averaged $1.1\text{--}1.2 \text{ m s}^{-1}$. Incorporating supplemental AWS observations and NDFD output decreased the MAE to 0.8 m s^{-1} at the 1-h lead time. Beyond the 10-h lead time, however, no improvements in CNDFD+ forecast accuracy (and, in fact, a slight worsening) occurred. Corresponding PC errors were nearly identical to those of the CNDFD+ for the 1–3-h lead times, after which an increasing loss of accuracy relative to CNDFD+ was demonstrated. In contrast to dewpoint, PC predictions of wind speed by the 5-h lead time became less accurate than the baseline CNDFD predictions.

a. CNDFD+ forecast skill

Alternatively, the above results can be translated into CNDFD+ percent improvements (i.e., MAE reduction; forecast skill). Figure 4 shows the improvements of adding AWS observations and supplemental NDFD forecasts to the calibrated NDFD forecast (i.e., skill of CNDFD+ predictions relative to corresponding CNDFD predictions). The figure plots each AWS station's skill

as a function of lead time for each parameter, with the black line an average over all tested sites. Several points can be gleaned from Fig. 4. As one might expect, an exponential decrease in skill with lead time occurs for all three parameters. Percentage improvements for temperature and dewpoint are comparable in magnitude, averaging 65% (0.65 translated as a skill score) at the 1-h lead time, exponentially decreasing to 20%–28% at the 6-h lead time, and 6%–7% by 12 h. Wind speed skill is markedly lower, averaging 22% at the 1-h lead time, and dropping <5% for forecasts beyond 5 h. Most stations show no CNDFD+ skill in wind speed beyond this time—a reflection of the high variability, and thus low predictability, of this variable.

Figure 5 presents CNDFD+ skill over reference PC forecasts. For the 1-h lead time, CNDFD+ forecast skill is tantamount to PC skill for dewpoint and wind speed, with a ~10% improvement for temperature. Regardless of parameter, CNDFD+ percent improvements gradually increase with lead time, consistent with Fig. 3, which revealed PC error increasing more rapidly than corresponding CNDFD+ error. By 6 h, CNDFD+ temperature improvements are 28%, and reach 36% by 12 h. Dewpoint and wind speed improvements are less, increasing to 16% and 12%, respectively, at the 6-h lead time, and slightly higher thereafter. Overall, Figs. 4 and 5 imply that PC remains a powerful forecast technique for the ultrashort-term (1–3 h) lead times, and that baseline CNDFD forecasts demonstrate value (i.e., have skill similar to CNDFD+ forecasts) for the longer lead times tested in this study. This result suggests that nonstandardized observations provide the greatest value relative

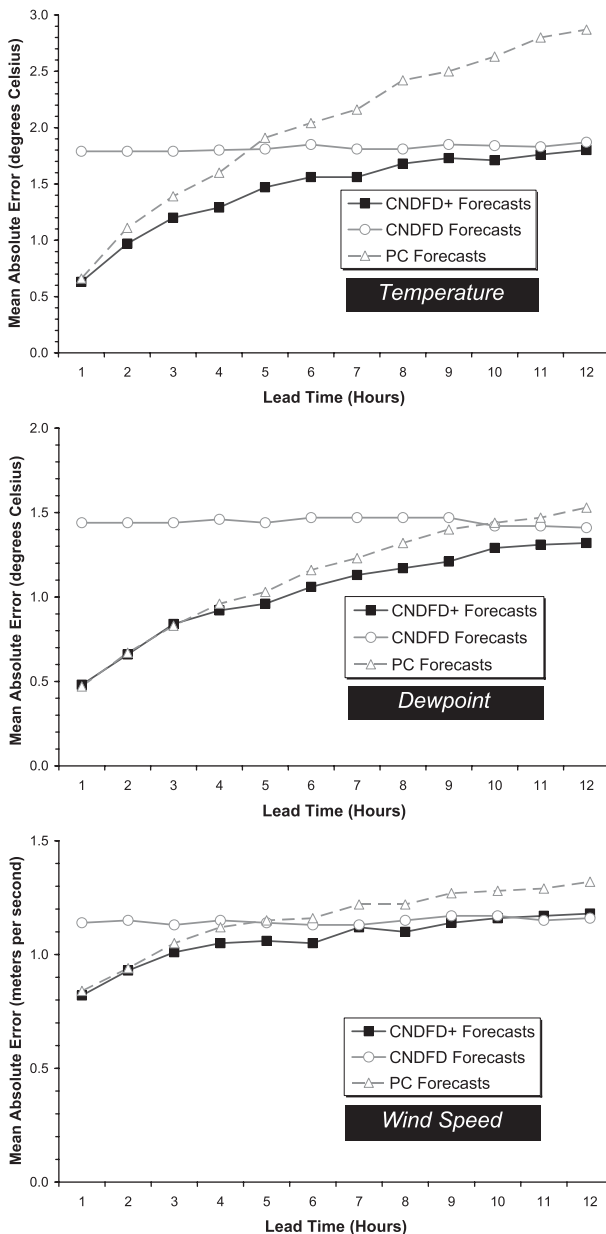


FIG. 3. MAE of (top) temperature at CHINM, (middle) dewpoint at SEASF, and (bottom) wind speed at BRTTN between verification and CNDFD+ forecasts (black solid line with squares), baseline CNDFD forecasts (gray solid line with circles), and PC forecasts (gray dashed line with triangles) as a function of lead time.

to CNDFD and PC [greatest relative accuracy (GRA) hereafter] for the middle lead times tested.

The lead time at which the GRA occurs (LT_{GRA} , hereafter) can be objectively determined by first superimposing the CNDFD and PC curves in Figs. 4 and 5, respectively, as constructed in Fig. 6. The next step is to highlight the segment of each curve indicating which of these two references for each lead time is more accurate

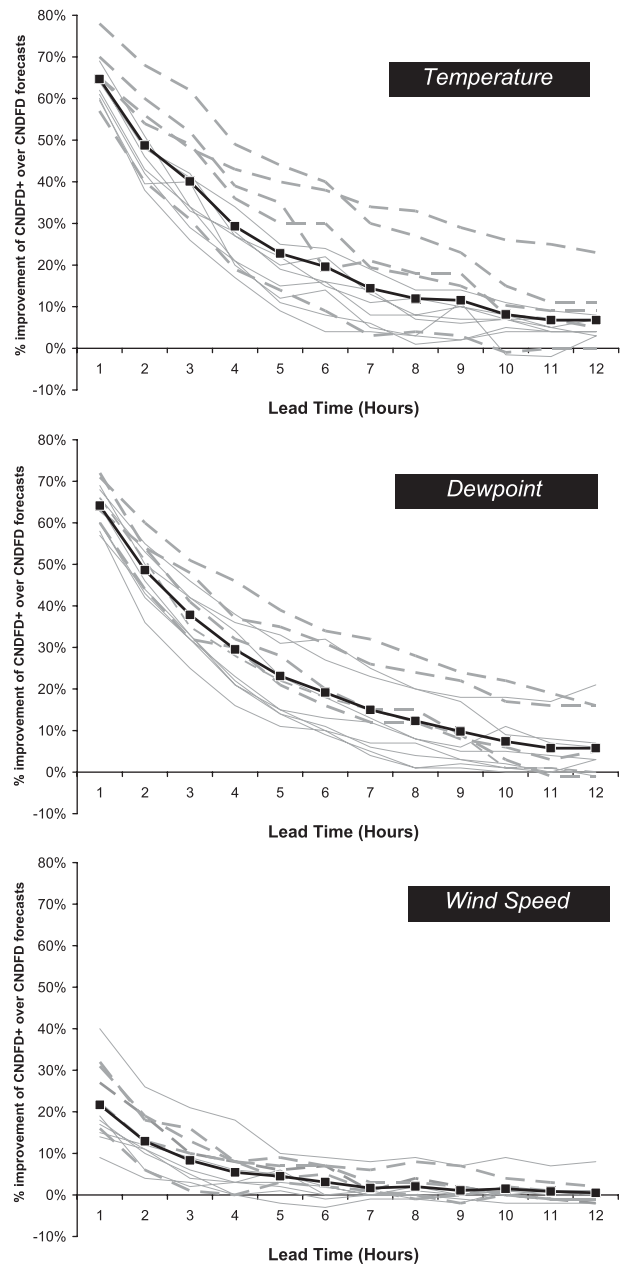


FIG. 4. Percent improvements of CNDFD+ forecasts over baseline CNDFD forecasts for (top) temperature, (middle) dewpoint, and (bottom) wind speed as a function of lead time for each AWS station. Tested stations with the highest elevations are dashed in gray. The black line with squares is an average over all tested AWS stations.

(in this context, possessing a skill score closer to 0 since 0 implies skill equivalent to CNDFD+). Figure 6 shows the blackened line for each parameter. The point at which the black line “jumps” from the PC to CNDFD curve—alternatively, when the black line reaches maximum skill—reveals the LT_{GRA} . As Fig. 6 shows, the

LT_{GRA} for all three parameters are in the middle lead times, ranging from 3–5 h for temperature and wind speed to 6–7 h for dewpoint. Note that percent improvements at each parameter's LT_{GRA} are in a different order—highest (23%) for temperature and lowest (6%) for wind speed.

An alternative method for assessing CNDFD+ forecast quality that integrates the suite of 1–12 lead times is to calculate the area under the black line (hatched area in Fig. 6). This “relative skill” (RS) can be expressed mathematically as an average:

$$RS = \frac{\sum_{lt=1}^{12} \text{Min}[\text{CNDFD}(lt), \text{PC}(lt)]}{12},$$

where lt is lead time and “Min[CNDFD(lt), PC(lt)]” is the minimum in percent improvement magnitudes of either CNDFD+ relative to CNDFD, or CNDFD+ relative to PC, as a function of lead time. The denominator allows the RS to be interpreted similarly to skill score: RS = 1 implies perfect value relative to the references; RS = 0 implies no value. Values of RS derived from Fig. 6 are 0.14 for temperature, 0.09 for dewpoint, and 0.03 for wind speed.

b. AWS siting variations

A station-by-station analysis of Figs. 4 and 5 reveals location-dependent differences in CNDFD+ forecast quality. Relative to CNDFD, CNDFD+ skill in forecasting temperature is greatest for MTQCR, BRTTN, Shenandoah National Park (LRAY1), and Wintergreen Mountain (WNTRE) (dashed, gray curves in Fig. 4). Table 1 shows that these four sites are a subset of the five highest elevation sites tested, with MELBA as the remaining site. Further exploration reveals higher-than-average CNDFD errors with these stations, implying the observations' effectiveness when adjusting CNDFD temperature forecasts in complex terrain. Spread in skill among stations is also evident with dewpoint and wind speed, with individual stations behaving differently. A weaker relationship exists between CNDFD+ performance and altitude for dewpoint, although skill of the higher-altitude stations remains near or above the 13-station average. No conclusive relationship exists for wind speed. A review of CNDFD+ performance (not shown) for stations located next to water [e.g., CHINM and Boardwalk Plaza Hotel (RBBPH)] also reveals only a weak correlation.

There are also AWS siting variations in CNDFD+ performance relative to PC, as Fig. 5 shows. The most striking pattern is that the majority of the higher-altitude

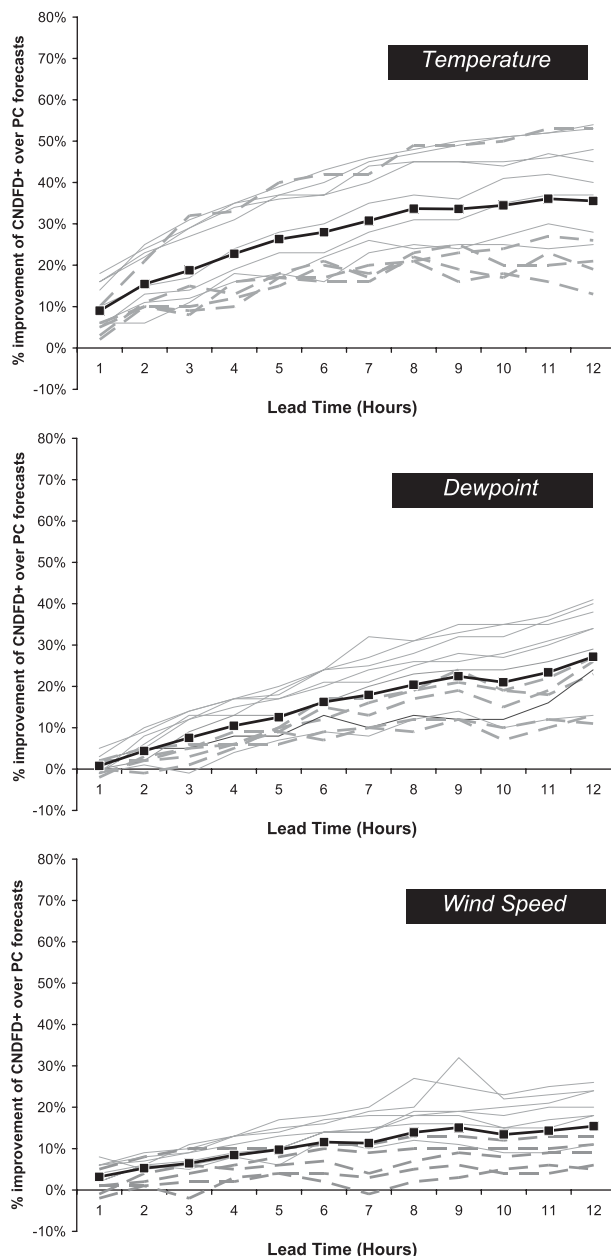


FIG. 5. As in Fig. 4, but percent improvements of CNDFD+ forecasts over reference PC forecasts.

stations, which demonstrated above-average performance relative to CNDFD, are *below* average relative to PC. This result suggests that the current observation of the forecast parameter at these sites is relatively more powerful than at the tested sites, and that the higher-altitude sites' auxiliary observations and NDFD forecasts are relatively less beneficial. This pattern seems to be most apparent for wind speed, in contrast to the lack of correlation between complex terrain sites and performance relative to CNDFD.

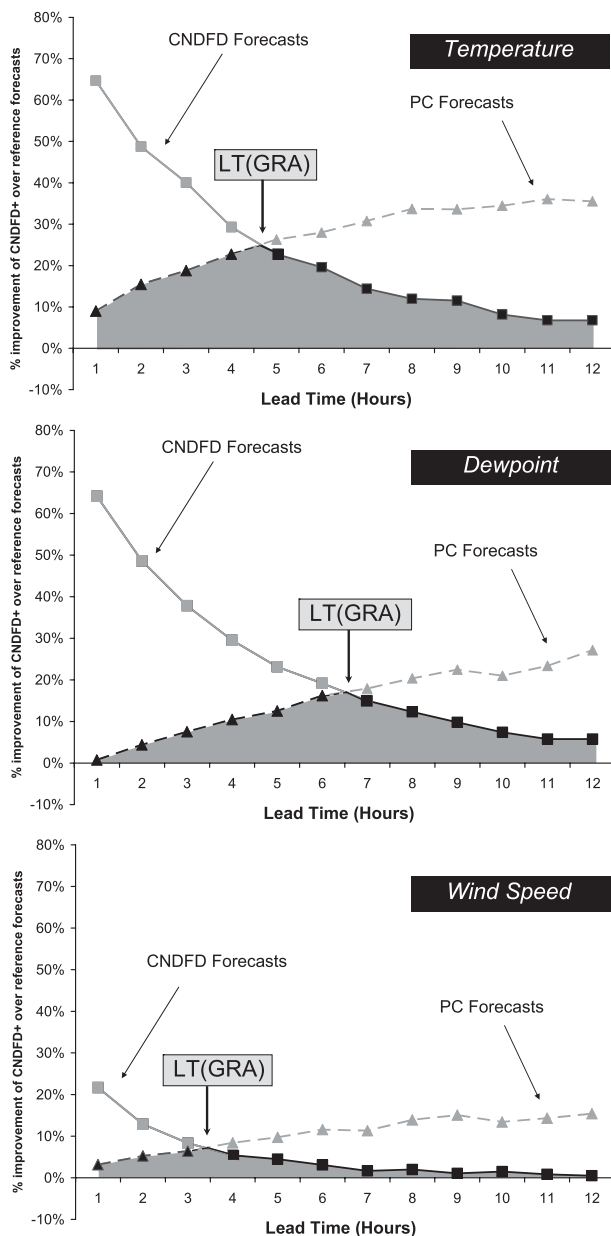


FIG. 6. Percent improvements of CNDFD+ forecasts over CNDFD forecasts (solid line with squares) and reference PC forecasts (dashed line with triangles) for (top) temperature, (middle) dewpoint, and (bottom) wind speed as a function of lead time. Data are averaged over all tested AWS stations. The blackened segment of each curve indicates whether CNDFD or PC is more accurate for that particular lead time. The lead time where the black line reaches maximum improvement is marked LT(GRA). The RS is derived from the shaded area.

Table 6 summarizes each AWS station's LT_{GRA} . Notable spread in values among locations is evident. Those stations with local effects generally have longer LT_{GRA} values (e.g., 10.5 h at MTQCR for temperature;

5.0-h RBBPH wind speed), while those with no apparent local effects typically have shorter LT_{GRA} values [e.g., 3.5-h Rohm and Haas (PHLRH) temperature; 4.5-h Peabody-Burns Elementary School (KPEAE) dewpoint]. Table 6 also shows the RS for each station for each parameter. The RS provides forecast quality information independent of the LT_{GRA} since the indices have some, but not perfect, correlation. For example, when compared with KPEAE (a site in Kansas with no apparent local effects), WNTRE (a site located atop Wintergreen Mountain in Virginia) consistently has a longer LT_{GRA} . On the other hand, RS values for dewpoint and wind speed at WNTRE are nearly identical to those of KPEAE.

6. Concluding remarks

This study's objective was to test the utility of a non-standardized surface observing network, such as the AWS, to improve 1–12-h calibrated NDFD (CNDFD) forecasts of energy parameters, including temperature, dewpoint, and wind speed. A sample of 13 AWS stations, located in varying terrain and proximity to water across the nation, was chosen to also explore relationships between forecast performance and AWS location. Forecasts were constructed by applying GAM, and allowing cubic smoothing splines to capture the nonlinearities between the predictand and its most significant predictors, ascertained using stepwise regression. Forecasts were then compared to two references: corresponding baseline CNDFD and PC forecasts. The main conclusions from this study follow:

- CNDFD+ forecast skill relative to PC increased slowly with lead time for all parameters. For dewpoint and wind speed, CNDFD+ forecast quality was comparable ($<5\%$) to PC for the ultrashort-term (1–3) lead times, increasing to $\sim 10\%$ – 15% by the 12-h lead time for wind speed and $\sim 25\%$ for dewpoint. CNDFD+ temperature predictions relative to PC were modestly more skillful: $\sim 10\%$ at the 1-h lead time, increasing to $\sim 30\%$ for lead times of 12 h.
- Averaged over all tested AWS stations, CNDFD+ temperature and dewpoint forecasts showed a 65% improvement in MAE over corresponding baseline CNDFD forecasts at the 1-h lead time. Forecast quality exponentially dropped to $\sim 20\%$ – 25% improvement for temperature and dewpoint by lead times of 6 h. By 12 h, CNDFD+ temperature and dewpoint forecasts were only marginally superior (6%–7% improvement) to CNDFD.
- AWS observations were only effective for improving CNDFD wind speed forecasts for the ultrashort-term

TABLE 6. LT_{GRA} (h) and RS for temperature, dewpoint, and wind speed of each tested AWS station. The bottom row represents the values combining all stations.

AWS identifier	Temperature— LT_{GRA} (RS)	Dewpoint— LT_{GRA} (RS)	Wind speed— LT_{GRA} (RS)
BRTTN	9.5 (0.134)	7.0 (0.056)	4.5 (0.026)
LWAY1	7.0 (0.128)	6.5 (0.071)	7.0 (0.017)
PHLRH	3.5 (0.110)	4.5 (0.066)	3.5 (0.033)
RBBPH	6.5 (0.108)	11.0 (0.104)	5.0 (0.078)
WIVBT	4.0 (0.115)	4.0 (0.048)	2.0 (0.015)
RCHSC	3.0 (0.108)	4.5 (0.103)	2.5 (0.026)
CHINM	4.5 (0.107)	5.5 (0.054)	3.0 (0.009)
BRRRD	3.5 (0.149)	5.5 (0.091)	2.0 (0.003)
WNTRE	7.0 (0.127)	6.5 (0.064)	4.5 (0.017)
MTQCR	10.5 (0.169)	>12.0 (0.079)	3.0 (0.003)
MELBA	3.0 (0.094)	>12.0 (0.080)	3.0 (0.063)
SEASF	6.0 (0.129)	10.0 (0.064)	— (—)
KPEAE	3.0 (0.099)	4.5 (0.062)	2.0 (0.023)
All stations	4.5 (0.140)	6.5 (0.090)	3.5 (0.030)

lead times as CNDFD+ wind speed improvements were notably less skillful than temperature or dewpoint: an $\sim 20\%$ improvement at the 1-h lead time, decreasing to $<5\%$ by 4 h.

- The greatest CNDFD+ accuracy relative to both CNDFD and PC was demonstrated for the middle lead times ($LT_{GRA} = 3\text{--}7$ h) tested in the study. The RS, a parameter designed to quantify the skill of CNDFD+ forecasts relative to CNDFD and PC, was highest (0.14) for temperature and lowest (0.03) for wind speed.
- Notable variations in CNDFD+ skill existed with respect to AWS location. The majority of the tested stations located in complex terrain generally showed forecast improvements relative to CNDFD greater than the 13-station average for temperature (and, to a lesser degree, dewpoint). Relative to PC, however, the majority of the sites located in complex terrain exhibited improvements *below* the 13-station average for all parameters.
- No conclusive relationship can be made between CNDFD+ forecast skill and the tested stations located near water, where local effects might influence the sensible weather.
- Siting variations in LT_{GRA} and RS values were also apparent, with some higher-altitude stations demonstrating a maximum skill relative to NDFD and PC beyond 8 h. Other sites, however, had wind speed RS values ~ 0.00 , revealing the lack of success of this technique for certain weather parameters and locations.

Overall, this study supports the potential benefits of nonstandardized surface networks and mesonetworks in constructing skillful short-term operational forecasting products (e.g., Hart et al. 2004; NDFD-derived gridded

MOS), with an emphasis here in adjusting calibrated NDFD forecasts. Based on a 1-yr data sample and 13 AWS sites, techniques employed in this work suggest that surface observations provide the greatest value for lead times of $\sim 3\text{--}7$ h. For shorter lead times, persistence climatology is a sensible strategy, while the impact of observations diminishes for longer lead times.

It is also suggested that this technique demonstrates added skill relative to the NDFD for locations in complex terrain, where local effects dominate and whose topography may not be well resolved in dynamic models. This result is particularly favorable for energy companies whose supply domain includes a moderate amount of complex and/or high terrain, such as *Energy West* in Montana and Wyoming. Conversely, the benefit will not be as great for those companies where the percentage of complex terrain, and thus population to impact load, is limited. Because some skill was shown for temperature and dewpoint, skillful heat index forecasts using nonstandardized observations may also be generated. On the other hand, the general lack of skill in forecasting wind speed may limit the observations' value during the winter in predicting wind chill.

It is reasonable to assume that the CNDFD+ improvement magnitudes ascertained in this study will decrease once NDFD forecasts are available hourly. Further changes in skill are anticipated once all WFOs update the NDFD grids every 3 h. With these modifications, additional and/or more precise NDFD predictors (e.g., NDFD₈, rather than the current NDFD₇₋₉, for an 8-h forecast) may be included in the forecast equations.

Although the AWS data used in the study are propriety, NRC (2008) includes a host of supplemental and publicly accessible networks and mesonetworks that are of comparable density, coverage, and quality to the AWS. These

networks add to the available sources for users wishing to incorporate observations in applications such as the one demonstrated here. Additional options are expected in the future as a national “network of networks” that relies on both proprietary and nonproprietary data evolves.

It is worth emphasizing that the use of AWS, or any supplemental network of surface observations, has limitations in adjusting a forecasting product such as the NDFD. The coverage of station observations is fractional when compared to the areal extent the NDFD offers across the United States. As such, the technique demonstrated here can only be applied where stations are available.

Thus, this work’s future direction includes exploring the observations’ radii of influence in improving nearby NDFD gridded forecasts. The present study can also be extended by assessing CNDP+ forecast improvement by including potential predictors from surrounding sites. Another opportunity would be to employ temporal and spatial correlation errors to improve short-term NDFD predictions of energy parameters. Finally, additional work spurred by the success of gridded MOS would test the nascent product’s utility when its output is complemented with that of the NDFD and observations from AWS and/or comparable surface observing networks.

Acknowledgments. The authors thank Jan Dutton, former Director of Professional Services, and Bill Callahan at AWS for granting access to their observational databases. We are also indebted to the NCDC for their Web-based availability of the NDFD archives. Moreover, the authors are grateful for the insightful and constructive comments by J. Paul Dallavalle and the anonymous reviewers of this work. We also thank Dean Iovino of the NWS WFO office in Philadelphia, PA, for offering his time and expertise. Finally, the authors recognize Sandra F. Mather, Professor Emeritus in the Department of Geology and Astronomy at West Chester University, for her contribution to the article’s publication. This study has been supported by a grant courtesy of AWS and the College of Arts and Sciences at West Chester University.

REFERENCES

- AWS, 2007: WeatherBug Weather Station Surface Observing System: User’s guide. WeatherBug Professional, 66 pp. [Available from AWS Convergence Technologies, Inc., 12410 Milestone Center Drive, Suite 300, Germantown, MD 20876.]
- , cited 2010: About us: WeatherBug overview. [Available online at <http://weather.weatherbug.com/about-us/about-us.html>.]
- Baggaley, D. G., and J. M. Hanesiak, 2005: An empirical blowing snow forecast technique for the Canadian Arctic and the Prairie Provinces. *Wea. Forecasting*, **20**, 51–62.
- Benjamin, S. G., B. E. Schwartz, and R. E. Cole, 1999: Accuracy of ACARS wind and temperature observations determined by collocation. *Wea. Forecasting*, **14**, 1032–1038.
- , and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.
- Bolzern, P., G. Fronza, and G. Brusasca, 1982: Temperature effects on the winter daily electric load. *J. Appl. Meteor.*, **21**, 241–243.
- Bondarenko, V., T. Ochotta, D. Saupe, and W. Wergen, 2007: The interaction between model resolution, observation resolution and observation density in data assimilation: A two-dimensional study. Preprints, *11th Symp. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, San Antonio, TX, Amer. Meteor. Soc., 5.19. [Available online at <http://ams.confex.com/ams/pdfpapers/117655.pdf>.]
- Buell, C. E., 1958: Meaning of combined climate and persistence forecasts. *J. Meteor.*, **15**, 564–565.
- Cosgrove, R. L., and B. Sfanos, 2004: Producing MOS snowfall amount forecasts from cooperative observer reports. Preprints, *20th Conf. on Weather Analysis and Forecasting*, Seattle, WA, Amer. Meteor. Soc., 6.3. [Available online at <http://ams.confex.com/ams/pdfpapers/69445.pdf>.]
- Dagostaro, V. J., W. A. Shaffer, M. J. Schenk, J. L. Gorline, and A. A. Taylor, 2004: A prototype verification system for examining NDFD forecasts. Preprints, *17th Conf. on Probability and Statistics in the Atmospheric Sciences*, Seattle, WA, Amer. Meteor. Soc., 2.7. [Available online at <http://ams.confex.com/ams/pdfpapers/73589.pdf>.]
- Daley, R., 1993: Estimating observation error statistics for atmospheric data assimilation. *Ann. Geophys.*, **11**, 634–647.
- Dallavalle, J. P., and B. Glahn, 2005: Toward a gridded MOS system. Preprints, *21st Conf. on Weather Analysis and Forecasting*, Washington, DC, Amer. Meteor. Soc., 13B.2. [Available online at <http://ams.confex.com/ams/pdfpapers/94998.pdf>.]
- , M. C. Erickson, and J. C. Maloney III, 2004: Model Output Statistics (MOS) guidance for short-range projections. Preprints, *20th Conf. on Weather Analysis and Forecasting*, Seattle, WA, Amer. Meteor. Soc., 6.1. [Available online at <http://ams.confex.com/ams/pdfpapers/73764.pdf>.]
- De Boor, C., 1978: *A Practical Guide to Splines*. Springer-Verlag, 392 pp.
- DiMego, G., 2006: WRF-NMM & GSI analysis to replace Eta model & 3DVar in NAM decision brief. National Centers for Environmental Prediction, 115 pp. [Available online at http://www.emc.ncep.noaa.gov/WRFinNAM/Update_WRF-NMM_replacing_Eta_in_NAM2.pdf.]
- Du, J., and Coauthors, 2009: NCEP short-range ensemble forecast (SREF) system upgrade in 2009. Preprints, *19th Conf. on Numerical Weather Prediction and 23rd Conf. on Weather Analysis and Forecasting*, Omaha, NE, Amer. Meteor. Soc., 4A.4.
- Efroymsen, M. A., 1960: *In Mathematical Methods for Digital Computers*. John Wiley & Sons, 191 pp.
- Fabbian, D., R. de Dear, and S. Lellyett, 2007: Application of artificial neural network forecasts to predict fog at Canberra International Airport. *Wea. Forecasting*, **22**, 372–381.
- Geerts, B., 2008: Dryline characteristics near Lubbock, Texas, based on radar and West Texas Mesonet data for May 2005 and May 2006. *Wea. Forecasting*, **23**, 392–406.
- Ghirardelli, J. E., 2005: An overview of the redeveloped Localized Aviation MOS Program (LAMP) for short-range forecasting.

- Preprints, *21st Conf. on Weather Analysis and Forecasting*, Washington, DC, Amer. Meteor. Soc., 13B.5. [Available online at <http://ams.confex.com/ams/pdfpapers/95038.pdf>.]
- , and B. Glahn, 2010: The Meteorological Development Laboratory's aviation weather prediction system. *Wea. Forecasting*, **25**, 1027–1051.
- Glahn, B., and D. P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.*, **84**, 195–201.
- , K. K. Gilbert, R. Cosgrove, D. Ruth, and K. Sheets, 2008: Gridded MOS guidance in the national digital guidance database. Preprints, *19th Conf. on Probability and Statistics*, New Orleans, LA, Amer. Meteor. Soc., 11.3. [Available online at <http://ams.confex.com/ams/pdfpapers/132526.pdf>.]
- , —, —, —, and —, 2009: The gridding of MOS. *Wea. Forecasting*, **24**, 520–529.
- Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.
- Grover-Kopec, E. K., and J. M. Fritsch, 2003: The impact of radar data on short-term forecasts of surface temperature, dewpoint depression, and wind speed. *Wea. Forecasting*, **18**, 1230–1241.
- Hart, K. A., W. J. Steenburgh, D. J. Onton, and A. J. Siffert, 2004: An evaluation of mesoscale-model-based Model Output Statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Wea. Forecasting*, **19**, 200–218.
- Hastie, T. J., and R. J. Tibshirani, 1990: *Generalized Additive Models*. Chapman and Hall, 335 pp.
- Helms, D., cited 2005: Citizen Weather Observer Program (CWOP): Weather station siting, performance, and data quality guide. CWOP Doc., 88 pp. [Available online at http://home.comcast.net/~dshelms/CWOP_Guide.pdf.]
- Hilliker, J. L., R. Vislocky, and J. M. Fritsch, 1999: An observations-based statistical system for warm-season hourly probabilistic forecasts of low ceiling at the San Francisco International Airport. *J. Appl. Meteor.*, **38**, 1692–1705.
- , G. S. Young, and J. M. Fritsch, 2007: An objective, statistical forecast system for short-term probabilistic forecasts of convection. *Natl. Wea. Dig.*, **31** (1), 9–23.
- Horel, J., and Coauthors, 2002: Mesowest: Cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83**, 211–225.
- Hughes, K. K., 2001: Development of MOS thunderstorm and severe thunderstorm forecast equations with multiple data sources. Preprints, *18th Conf. on Weather Analysis and Forecasting*, Fort Lauderdale, FL, Amer. Meteor. Soc., 191–195.
- Illton, B. G., J. B. Basara, D. K. Fisher, R. Elliott, C. A. Fiebrich, K. C. Crawford, K. Humes, and E. Hunt, 2008: Mesoscale monitoring of soil moisture across a statewide network. *J. Atmos. Oceanic Technol.*, **25**, 167–181.
- Janjić, T., and S. E. Cohn, 2006: Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134**, 2900–2915.
- Leyton, S. M., and J. M. Fritsch, 2003: Short-term probabilistic forecasts of ceiling and visibility utilizing high-density surface weather observations. *Wea. Forecasting*, **18**, 891–902.
- Liu, Z., and F. Rabier, 2002: The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J. Roy. Meteor. Soc.*, **128**, 1367–1386.
- Ludwig, F. L., J. Horel, and C. D. Whiteman, 2004: Using EOF analysis to identify surface wind patterns in mountain valleys. *J. Appl. Meteor.*, **43**, 969–983.
- Martin, F. L., J. R. Borsting, F. J. Steckbeck, and A. H. Manhard Jr., 1963: Statistical prediction methods for North American anticyclones. *J. Appl. Meteor.*, **2**, 508–516.
- Miller, P. A., M. Barth, L. Benjamin, D. Helms, M. Campbell, J. Facundo, and J. O'Sullivan, 2005: The Meteorological Assimilation and Data Ingest System (MADIS)—Providing value-added observations to the meteorological community. Preprints, *21st Conf. on Weather Analysis and Forecasting*, Washington, DC, Amer. Meteor. Soc., P1.95.
- Mollner, M., 2005: Notes on the National Digital Forecast Database verification. Western Region Tech. Attachment 05-02, 8 pp. [Available online at <http://www.wr.noaa.gov/wrh/05TAs/ta0502.pdf>.]
- Murphy, A., 1992: Climatology, persistence, and their linear combination as standards of reference in skill scores. *Wea. Forecasting*, **7**, 692–698.
- Myrick, D. T., and J. D. Horel, 2006: Verification of surface temperature forecasts from the National Digital Forecast Database over the western United States. *Wea. Forecasting*, **21**, 869–892.
- , and —, 2008: Sensitivity of surface analyses over the western United States to RAWS observations. *Wea. Forecasting*, **23**, 145–158.
- Nadolski, V. L., 1998: Automated Surface Observing System. ASOS User's Guide, National Oceanic and Atmospheric Administration, Department of Commerce, 56 pp. [Available online at <http://www.nws.noaa.gov/asos/pdfs/aum-toc.pdf>.]
- National Research Council, 2008: *Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks*. National Academy Press, 250 pp.
- Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman, 1996: *Applied Linear Statistical Models*. 4th ed. Irwin Professional Publishing, 1408 pp.
- NWS, 1993: Quality control incoming data. Technique Specification Package 88-21-R1 for AWIPS-90, National Oceanic and Atmospheric Administration, Department of Commerce.
- , cited 2006: National Weather Service ASOS sites. [Available online at http://www.nws.noaa.gov/mirs/public/prods/maps/map_images/national-maps/asos_us.jpg.]
- Quixote Transportation Technologies, Inc., 2010: RWIS: Road/Runway Weather Information Systems, 2 pp. [Available online at http://www.qttinc.com/pdf/lit-center/RWIS%20Platforms_8-27-09.pdf.]
- Robinson, P. J., 1997: Modeling utility load and temperature relationships for use with long-lead forecasts. *J. Appl. Meteor.*, **36**, 591–598.
- Rothe, J. P., A. K. Wadhwani, and S. Wadhwani, 2009: Short term load forecasting using multi parameter regression. *Int. J. Comput. Sci. Inf. Secur.*, **6** (2), 303–306.
- Ruth, D. P., B. Glahn, V. Dagostaro, and K. Gilbert, 2009: The performance of MOS in the digital age. *Wea. Forecasting*, **24**, 504–519.
- Schmeits, M. J., K. J. Kok, D. H. P. Vogelesang, and R. M. van Westrhenen, 2008: Probabilistic forecasts of (severe) thunderstorms for the purpose of issuing a weather alarm in the Netherlands. *Wea. Forecasting*, **23**, 1253–1267.
- Serreze, M., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty, 1999: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resour. Res.*, **35**, 2145–2160.

- S-PLUS, 2001: S-PLUS 6 for Windows: User's guide. Insightful Corporation, 688 pp.
- Teisberg, T. J., R. F. Weiher, and A. Khotanzad, 2005: The economic value of temperature forecasts in electricity generation. *Bull. Amer. Meteor. Soc.*, **86**, 1765–1771.
- Valor, E., V. Meneu, and V. Caselles, 2001: Daily air temperature and electricity load in Spain. *J. Appl. Meteor.*, **40**, 1413–1421.
- Vislocky, R. L., and J. M. Fritsch, 1995: Generalized additive models versus linear regression in generating probabilistic MOS forecasts of aviation weather parameters. *Wea. Forecasting*, **10**, 669–680.
- , and —, 1997: An automated, observations-based system for short-term prediction of ceiling and visibility. *Wea. Forecasting*, **12**, 31–43.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 627 pp.
- Zachariassen, J., K. Zeller, N. Nikolov, and T. McClelland, 2003: A review of the Forest Service Remote Automated Weather Station (RAWS) network. General Tech. Rep. RMRS-GTR-119, Rocky Mountain Research Station, U.S. Forest Service, Fort Collins, CO, 153 pp. [Available online at http://www.fs.fed.us/rm/pubs/rmrs_gtr119.pdf.]

Copyright of Journal of Applied Meteorology & Climatology is the property of American Meteorological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.