

The New England High-Resolution Temperature Program

BY DAVID J. STENSRUD, NUSRAT YUSSOUF, MICHAEL E. BALDWIN, JEFFERY T. MCQUEEN, JUN DU, BINBIN ZHOU, BRAD FERRIER, GEOFFREY MANIKIN, F. MARTIN RALPH, JAMES M. WILCZAK, ALLEN B. WHITE, IRINA DJLALOVA, JIAN-WEN BAO, ROBERT J. ZAMORA, STANLEY G. BENJAMIN, PATRICIA A. MILLER, TRACY LORRAINE SMITH, TANYA SMIRNOVA, AND MICHAEL F. BARTH

A collaborative research project brings together the operational and research components of NOAA to improve 2-m temperature forecasts for the New England region and to engage all sectors of the weather and climate enterprise.

The New England High-Resolution Temperature Program (NEH RTP) began in early 2002 through a Congressional budget request instructing the National Oceanic and Atmospheric Administration (NOAA) to design and implement a research and development program to improve the accuracy of summertime temperature forecasts in New England.

The intent was that improved temperature forecasts would lead to more accurate forecasts of electricity demand by utility companies, helping these companies to reduce costs and conserve energy, and then pass along these savings to consumers.

In response to this request, NOAA put together a research and development team that spans and integrates the research and operational components of the agency. The operational component is led by the Environmental Modeling Center (EMC) of the National Centers for Environmental Prediction (NCEP), with additional input and assessment provided by National Weather Service (NWS) forecasters throughout the New England region. The research component is represented by the Environmental Technology Laboratory (ETL), the Forecast Systems Laboratory (FSL), and the National Severe Storms Laboratory (NSSL). These four main groups have a wealth of experience using, developing, and assessing observational systems and numerical weather prediction models.

Initial discussions with representatives from several electric utilities indicated that improvements in the day-two forecasts (i.e., today's forecast for tomorrow) would be most helpful to their industry, and that during the summertime the most important atmospheric variables influencing power demand are

AFFILIATIONS: STENSRUD, YUSSOUF, AND BALDWIN—NOAA/National Severe Storms Laboratory, Norman, Oklahoma; MCQUEEN, DU, ZHOU, FERRIER, AND MANIKIN—NOAA/NWS/NCEP/Environmental Modeling Center, Camp Springs, Maryland; RALPH, WILCZAK, WHITE, DJLALOVA, BAO, AND ZAMORA—NOAA/Environmental Technology Laboratory, Boulder, Colorado; BENJAMIN, MILLER, SMITH, SMIRNOVA, AND BARTH—NOAA/Forecast Systems Laboratory, Boulder, Colorado

CORRESPONDING AUTHOR: Dr. David J. Stensrud, NOAA/National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069

E-mail: david.stensrud@noaa.gov

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-87-4-491

In final form 16 January 2006
©2006 American Meteorological Society

2-m temperature, 2-m dewpoint temperature, and 10-m winds (in general order of importance). Thus, our early planning meetings focused largely on how to leverage current capabilities and expertise to improve forecasts of these three near-surface variables.

The team decided upon a four-pronged approach to improve the near-surface forecasts in New England that includes 1) improved surface and boundary layer observations for model initialization, 2) special observations for the assessment and improvement of model physical process parameterization schemes, 3) use of model forecast ensemble data to improve upon the present operational forecasts for the three near-surface variables, and 4) transfer of knowledge gained from NEH RTP to both the commercial weather services and the electric power utilities. Because the focus of the project is on summertime predictions, special data collection activities occurred during the summers of 2002 through 2004. Our accomplishments during this program in each of the four identified areas now are highlighted.

ENHANCED OBSERVATIONS. The first major goal of NEH RTP was to enhance the existing observational network in and around the northeastern part of the country. This enhancement was achieved by moving in special NOAA instrumentation for the duration of NEH RTP, and by collecting and utilizing the many other-agency observations already available in the area. Owing to the goal of improving near-surface forecasts, boundary layer profilers and improved surface sites were the main focus for enhancements.

Boundary layer wind profilers and Radio Acoustic Sounding Systems (RASS) were installed and operated across New England during each of the three summers, with additional profilers deployed in neighboring regions (Fig. 1). At least eight profilers were deployed each summer, with the data from these instruments also being used for NOAA's New England air quality studies in 2002 and 2004. During the summer of 2004, other special instrumentation was deployed at a central supersite near Plymouth, Massachusetts, that provided more detailed information on each of the terms within the surface energy budget (Table 1). Observations also were acquired from nine other-agency profilers and RASS whose data were neither yet centrally collected nor operationally used (Fig. 1).

The available surface measurements of temperature were enhanced by automating the Cooperative Observer Program (COOP) sites to provide a more uniform density of surface observations (Fig. 2). Approximately 70 new COOP sites were operational by the end of 2004. In addition, pressure, temperature,

humidity, wind, and precipitation observations from more than 1000 mesonet sites in the Northeast (Fig. 2) were collected, integrated with other available surface observations (including the newly automated COOP observations), quality controlled, and distributed to NWS forecasters, NCEP, NEH RTP participants, and the greater meteorological community. Data ingest, integration, quality control, and distribution were achieved using the Meteorological Assimilation Data Ingest System (MADIS; MacDermaid et al. 2005; Miller et al. 2005) in conjunction with the Cooperative Agency Profiler (CAP) program.

While not all of the special observational capabilities are still in place, the majority of the new data streams acquired through NEH RTP now represent routine observations that are available via MADIS and are provided to the NWS for use in forecast offices and for generating operational model initial conditions. The boundary layer profiler and some of the new me-

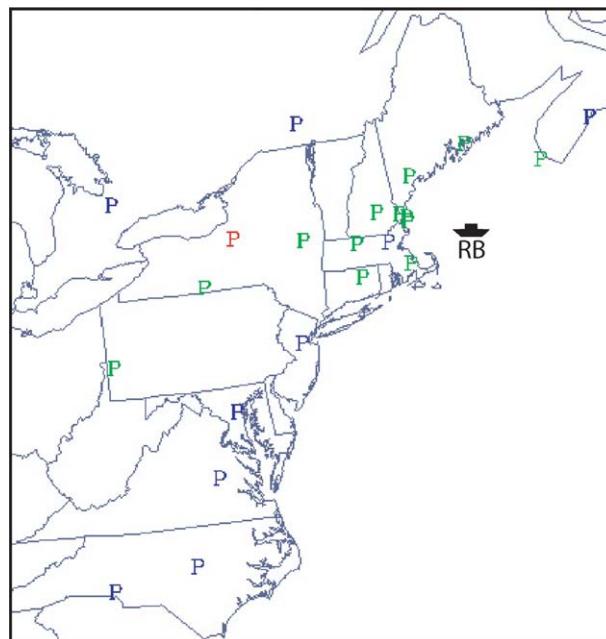


FIG. 1. Locations of the profiler sites for NEH RTP. NOAA profilers are indicated in red and green, with green indicating profilers moved into the New England area. The ETL special profiler installations included an installation on the NOAA research vessel the *Ron Brown* (RB) that cruised the coastal waters during the summers of 2002 and 2004. The locations of other-agency profilers collected for NEH RTP are indicated in blue. Agencies contributing profiler data included Environment Canada, the North Carolina Department of Environment and Natural Resources, McGill University, Maryland Department of the Environment, Virginia Department of Environmental Quality, Rutgers University, and the Massachusetts Department of Environmental Protection.

TABLE 1. Observing systems and measurements taken during the 2004 NEH RTP summer experiment at the Plymouth, MA, supersite. Sampling resolution also is indicated.

Parameter	Method	Sampling resolution (space and time)
Wind profile	915-MHz Doppler wind profiler	60–100 m, hourly
Temperature profile	Radio Acoustic Sounding System	100 m, hourly
Temperature profile	60-GHZ radiometer	2–200 m, 15 min
Cloud profile	S-band radar	45 m, 5 min
Integrated water vapor and cloud liquid water	Radiometer	10 min
CBL depth	915-MHz profiler	60–100 m, hourly
Nocturnal boundary layer structure	Bistatic backscatter sodar	15 m, 5 min
Surface heat and momentum fluxes	Applied Technologies, Inc. (ATI), sonic anemometer	20 m, 30 min
Humidity and CO fluxes	LICOR 7500	20 m, 30 min
Surface wind	RM Young wind monitor	10 m, 20 m, 2 min
Pressure	Vaisala analog pressure probe	1 m, 2 min
Temperature, RH	Vaisala HMP45C	2 m, 20 m, 2 min
Solar radiation	Kipp & Zonen pyranometer	2 m, 2 min
Net radiation	Radiation and Energy Balance Systems (REBS) net radiometer	2 m, 2 min
Rainfall	Texas Electronics tipping bucket	1 m, 2 min
Aerosol optical depth	Sun photometer	1 m, hourly
Four-stream radiation	Eppley	20 m, hourly
Direct/diffuse solar	Eppley	1 m, hourly
Soil temperature	CSI-107	5, 10, 15, and 60 cm, hourly
Soil moisture	CSI-660	10 and 60 cm, hourly
Ground heat flux	HFT3	2 cm, hourly

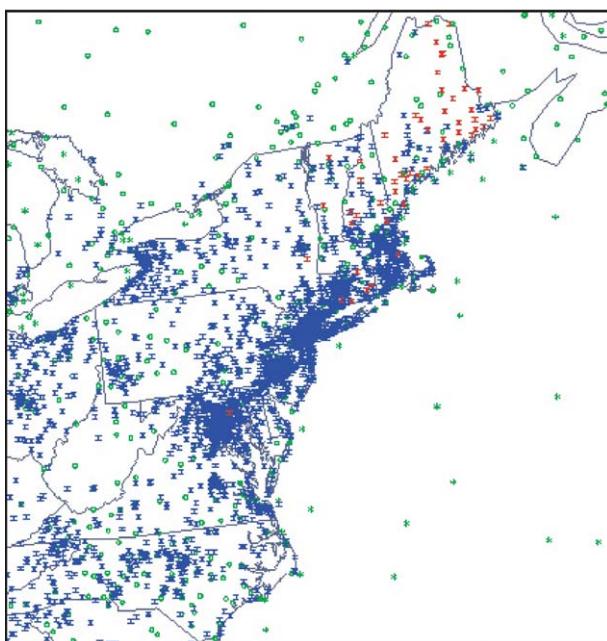


FIG. 2. Station locations for surface observations collected for the NEH RTP. Meteorological Terminal Air Report (METAR) and maritime observations are shown in green, modernized COOP stations in red, and other-agency mesonet observations in blue. Organizations contributing mesonet observations for NEH RTP included AWS Convergence Technologies, Inc., the Citizen Weather Observer Program (CWOP), WeatherforYou.com, Anything Weather, the Gulf of Maine Ocean Observing System, and the U.S. Army's Aberdeen Proving Grounds. In addition, the National Ocean Service provided observations from its Physical Oceanographic Real-Time System (PORTS), the Interagency Fire Center contributed Remote Automated Weather Stations (RAWS) observations, FSL contributed observations from its Ground-Based GPS Meteorology (GPS-Met) network, and the Department of Agriculture contributed observations from its Soil Climate Analysis Network (SCAN).

sonet data were used to improve the initial conditions of the rapid update cycle (RUC) model (Benjamin et al. 2004) both during and after the summer NEHRTP experiments. This outcome represents a substantial improvement in the ability of NOAA to sample the near-surface conditions in New England, thanks to the willingness of the private sector and other agencies to share their data for public benefit.

MODEL PARAMETERIZATION SCHEME EVALUATION. Each summer, special forecasts from operational and research numerical weather prediction models were run out to 48 h from a 1200 UTC initial time, and the data were archived at a central location. These forecasts included the 15 members of the 32-km short-range ensemble forecast (SREF) system at NCEP (Du et al. 2004), the operational Eta Model (Black 1994), the Global Forecast System model (Kalnay et al. 1990), several different versions of the RUC model using grid spacings of 20 km or less (Benjamin et al. 2004), the Nonhydrostatic Mesoscale Model (NMM; Janjic et al. 2001), a modified 22-km version of the Eta Model (Kain et al. 2001), the Weather Research and Forecast (WRF) model (Klemp 2004), and, in 2002, the fifth-generation Pennsylvania State University (PSU)–National Center

for Atmospheric Research (NCAR) Mesoscale Model (Dudhia 1993). These model forecasts varied not only in the model used and its grid spacing, but also in their physical process parameterization schemes and initial and boundary conditions. Each summer, anywhere from 24 to 31 forecasts were provided in real time with forecast data saved at 3-h intervals. These forecast data were then used to compare with observations in order to determine if any persistent biases exist in the forecast data that could lead to errors in near-surface conditions. The performance of each of the forecast models at each of the profiler sites, and at eight surface meteorological sites that are routinely used by energy providers to forecast energy loads in New England, were tracked in real time throughout each of the NEHRTP field campaigns, and were made available to forecasters through a NOAA Web page.

One of the principal findings of the analysis of the special observations taken during NEHRTP was the recognition that near-surface temperature errors in most of the models were strongly correlated to errors in the model-predicted radiation fields. As an example, the operational Eta Model and the NMM both demonstrated a diurnal variation in the 2-m temperature bias errors, with too warm daytime

temperatures and too cold nighttime temperatures (Fig. 3). Similar errors were found in many of the other models as well (not shown). Measurements of the solar radiation show that during the daytime hours these two models have too much solar radiation, reaching the surface during the period of too warm 2-m temperatures. Conversely, measurements of downward longwave radiation show that during the nighttime hours the models have too little downward longwave radiation during the period of too cold 2-m temperatures.

These results indicate that errors in the radiation parameterization as well as the cloud parameterization affect the accuracy of the forecasted radiation. To better understand the source of the radiation errors and their impact on surface temperatures through the surface energy balance, data from an observational “supersite” were collected during the summer of 2004 (see Table 1).

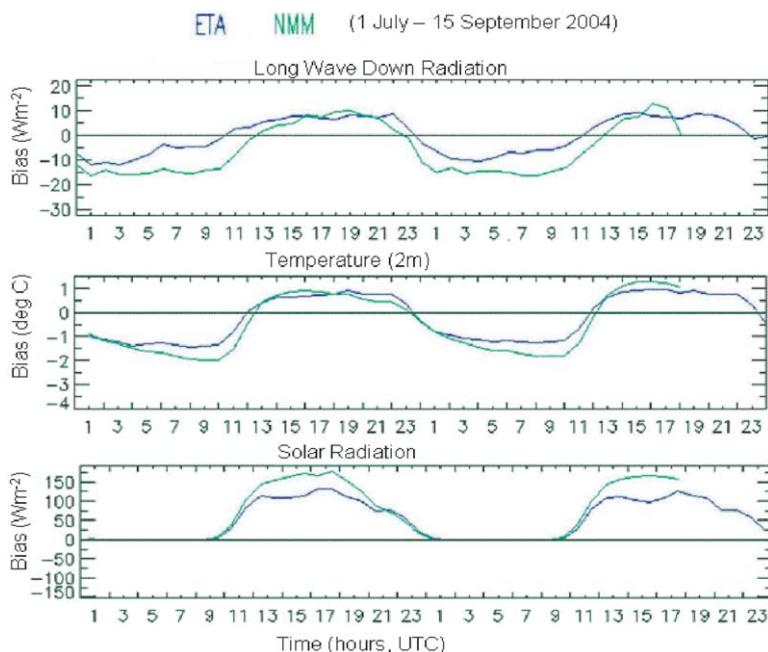


FIG. 3. Plots of bias (model – observed) in (top) downward longwave radiation, (middle) 2-m temperature, and (bottom) solar radiation versus time (UTC) over the 48-h forecast time and averaged over 77 days. Eta Model bias is in blue and NMM bias is in green. The downward longwave radiation was only observed at a single site (Plymouth, MA), whereas the solar radiation and 2-m temperature are averages from six sites.

These measurements are currently being analyzed to identify precisely those atmospheric conditions that lead to the radiation biases, to isolate the errors in the radiation and cloud parameterizations that are responsible for the radiation errors, and to evaluate the effects of these radiation errors on the surface energy balance and surface temperature forecasts. For example, one potential source for the observed solar radiation bias errors is the attenuation due to aerosols, because this effect typically is not accounted for in weather forecast models. Aerosol optical depth measurements made during NEH RTP already have been compared with observed and predicted solar radiation from the Eta Model (Zamora et al. 2005). Results indicate that for each 0.1 increase in aerosol optical depth, the observed solar radiation decreases by $\sim 12 \text{ W m}^{-2}$ (Fig. 4). As a result, solar noon radiance errors on days with high aerosol loading can easily reach $60\text{--}80 \text{ W m}^{-2}$. Applying a simple analytic model, Zamora et al. (2005) found that this magnitude of radiation error can produce a surface skin temperature error on the order of 1 K. Once the error sources for the radiation and cloud parameterizations are identified, improvements to the radiation and cloud parameterizations will then be tested, with the expectation that these will lead to improved forecasts of near-surface conditions.

ENSEMBLE BIAS CORRECTION SYSTEM.

One of the benefits of having an ensemble of models available routinely is that one can investigate new approaches for postprocessing the model data to improve forecasts without waiting for improvements in the model physical process schemes. To evaluate the ability of an ensemble forecast system to provide improved forecasts of near-surface conditions, a simple bias correction scheme was developed (Stensrud and Yussouf 2003, 2005). This bias-corrected ensemble (BCE) approach uses the past 12 days of surface and model data to remove the bias individually from each model and for each station location and forecast time. Once the bias has been removed from each model, the resulting BCE is compared against observations and model output statistics (MOS; Glahn and Lowry 1972;

Jacks et al. 1990) from the Eta Model and the Aviation (AVN) version of the Global Forecast System.

Results indicated that the BCE mean provided more accurate predictions of 2-m temperature and dewpoint temperature at all forecast times out to 48 h (Fig. 5)

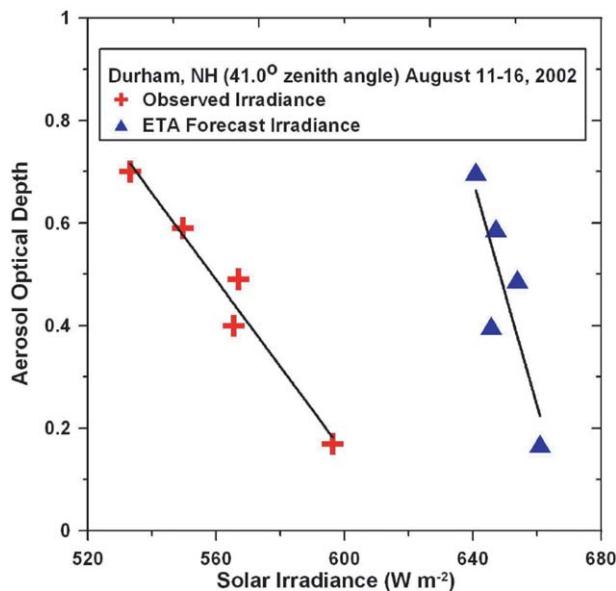


FIG. 4. Correlation between aerosol optical depth and the observed (crosses) and Eta Model (triangles) solar irradiances for a zenith angle of 41° , measured on five different cloud-free days.

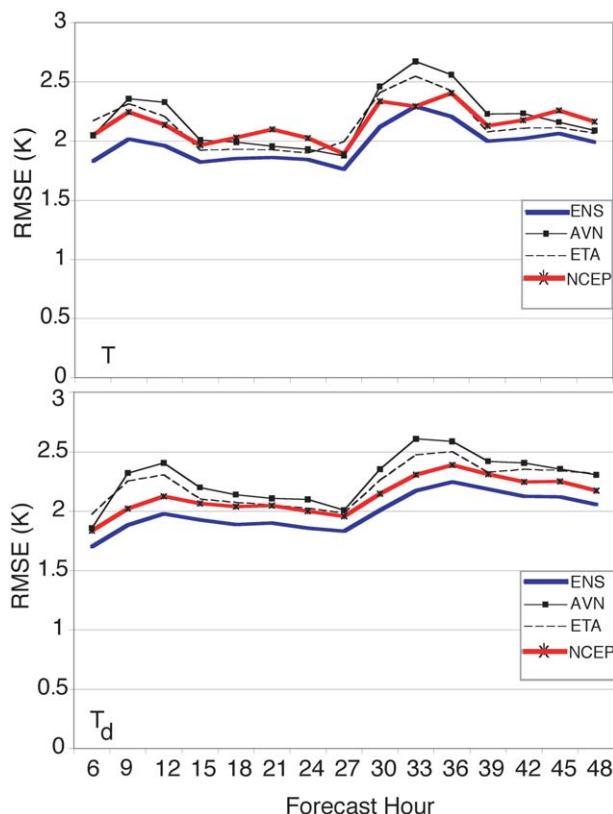


FIG. 5. Values of root-mean-square error (K) plotted as a function of forecast hour for (top) 2-m temperature and (bottom) 2-m dewpoint temperature from the full 31 member BCE (blue), the NCEP-only BCE (red), and the AVN (solid black line) and Eta (dashed line) MOS. Results are calculated at 1258 station locations for both the ensemble and AVN and Eta MOS data (after Stensrud and Yussouf 2005).

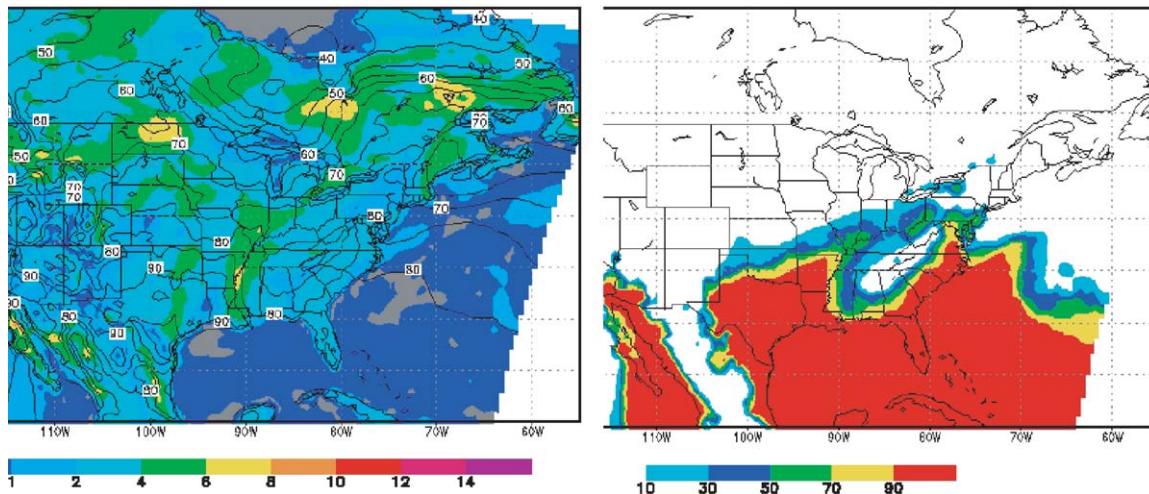


FIG. 6. Example of 32-km NCEP SREF 48-h forecasts valid at 0000 UTC 17 Sep 2004 showing (left) 2-m temperature ensemble mean and spread and (right) probability of the 2-m temperature–humidity index exceeding 75°F.

and provided equally accurate predictions of wind speed (not shown). The inclusion of additional models in the ensemble improved the BCE forecasts beyond that found when using the bias correction approach on only the operational model forecasts at NCEP (Fig. 5). In addition, the ensemble also provided information on forecast uncertainty and probabilities as illustrated by the NCEP 32-km short-range ensemble forecasting system (Fig. 6). The BCE probabilities were evaluated for several different threshold temperature values, and the results indicated that the probabilities were very reliable and skillful (Fig. 7). Simple cost–loss calculations (Katz and Murphy 1997) suggested that the BCE probabilities could lead to energy savings by utilities

if these BCE probability forecasts were used instead of the deterministic MOS forecasts. Thus, techniques for the postprocessing of model data developed by NEHRTP can provide improved 2-m temperature and 2-m dewpoint temperature forecasts over New England as requested by Congress. These techniques are easy to implement and do not require a large computing infrastructure, such that they are simple to migrate to new systems and respond quickly to forecast model changes. It is clear that the simple postprocessing of the ensemble data adds value and should be a routine part of any numerical weather prediction model forecast system.

KNOWLEDGE DISSEMINATION. The final goal of NEHRTP was to disseminate the knowledge and techniques developed during this project that were aimed at assisting electric utilities. Routinely providing tailored products for the energy sector is not part of NOAA’s operational mission, so a workshop was held in November 2004 to share our results with both the commercial weather services and

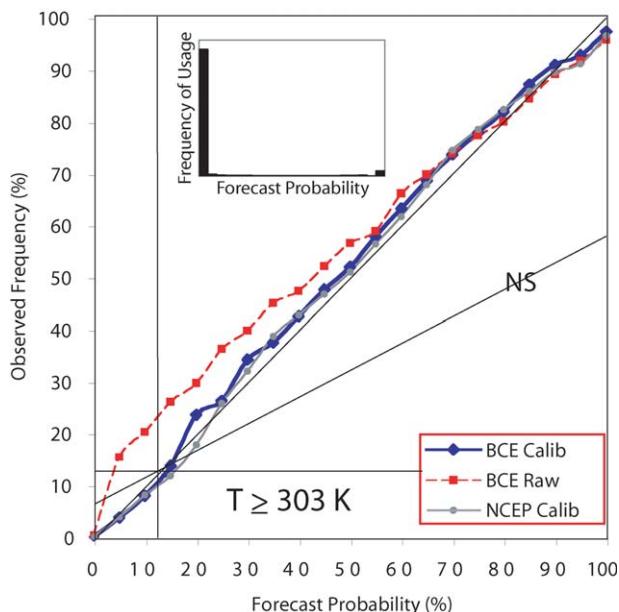


FIG. 7. Attribute diagram for the BCE forecasts of 2-m temperature equal to or exceeding 303 K, with calibration for the full ensemble (blue line) and the NCEP-only ensemble (gray line). Results from the uncalibrated full-ensemble data also are shown (red dashed line). The inset histogram indicates the frequency of usage of each 5% interval forecast probability category for the uncalibrated (raw) ensemble. Horizontal line indicates the frequency of the event in the observed dataset, and the diagonal line is the no-skill (NS) line (after Stensrud and Yussouf 2005).

the electric utilities as reported in Stensrud (2006). Feedback from this workshop has generally been very positive. In addition, results will be shared in the formal scientific literature as the research studies are completed (e.g., Stensrud and Yussouf 2003, 2005; Yussouf et al. 2004), and it is hoped that this article also will help to disseminate the results found during the NEHRTP campaigns of 2002 through 2004.

CONCLUSIONS. NEHRTP was developed in response to a request from Congress to improve temperature forecasting in New England with the specific intent of helping electric utilities, but it also has led to a number of unexpected benefits. The interaction between the NOAA groups has been exceptionally positive, and illustrates the tremendous benefits that can accrue when the research and operational groups collaborate toward common goals at an individual scientist level. A greater use of the available surface and boundary layer observations in New England is now a reality, thanks to cooperation from the private sector and other government agencies with program participants. In addition, the observations and model forecasts collected and analyzed during the three summer experimental periods are yielding insight into errors within the model physical process parameterization schemes. These analyses likely will lead to modifications in these schemes to improve operational forecasts of near-surface variables. Finally, a simple bias-corrected ensemble postprocessing technique already can yield improved forecasts of 2-m temperature and 2-m dewpoint temperature over New England and the rest of the United States when compared to the operational MOS forecasts. This technique is computationally inexpensive and easy to implement, and it is believed that many end users of weather forecast information could benefit from a probabilistic approach to weather forecasting. It remains to be seen if all of the benefits found from NEHRTP will be used to provide improved forecasts of near-surface variables across the United States, but it is hoped that one result of NEHRTP is that a new era of public-private partnership in the weather and climate enterprise has begun.

ACKNOWLEDGMENTS. We greatly appreciate the assistance of the many NWS forecasters in the New England region who provided input and support during NEHRTP. We also appreciate the project support provided by Lee Stang, David Green, and Daniel Melendez of the NWS. In addition, the many government agencies and private sector firms who willingly provided surface and boundary layer data to NOAA for public benefit and use are thanked and

have our sincere appreciation and respect. Funding for NEHRTP was provided through a Congressional request initiated by Senator Judd Gregg of New Hampshire.

REFERENCES

- Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132**, 473–494.
- Black, T. L., 1994: The new NMC mesoscale Eta Model: Description and forecast examples. *Wea. Forecasting*, **9**, 265–278.
- Du, J. J., and Coauthors, 2004: The NOAA/NWS/NCEP short-range ensemble forecast (SREF) system: Evaluation of an initial condition vs multiple model physics ensemble approach. Preprints, *16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., CD-ROM, 21.3.
- Dudhia, J., 1993: A nonhydrostatic version of the Penn State-NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**, 1493–1513.
- 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.
- Jacks, E., J. B. Bower, V. J. Dagostaro, J. P. Dallavalle, M. C. Erickson, and J. C. Su, 1990: New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and surface wind. *Wea. Forecasting*, **5**, 128–138.
- Janjic, Z. I., J. P. Gerrity Jr., and S. Nickovic, 2001: An alternative approach to nonhydrostatic modeling. *Mon. Wea. Rev.*, **129**, 1164–1178.
- Kain, J. S., M. E. Baldwin, P. Janish, and S. J. Weiss, 2001: Utilizing the Eta Model with two different convective parameterizations to predict convective initiation and evolution at the SPC. Preprints, *Ninth Conf. on Mesoscale Processes*, Ft. Lauderdale, FL, Amer. Meteor. Soc., 91–95.
- Kalnay E., M. Kanamitsu, and W. E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. *Bull. Amer. Meteor. Soc.*, **71**, 1410–1428.
- Katz, R. W., and A. H. Murphy, 1997: *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, 222 pp.
- Klemp, J., 2004: Next-generation mesoscale modeling: A technical overview of WRF. Preprints, *20th Conf. on Weather Analysis and Forecasting*, Seattle, WA, Amer. Meteor. Soc., CD-ROM, 11.2.
- MacDermaid, C., R. C. Lipschutz, P. Hildreth, R. A. Ryan, A. B. Stanley, M. F. Barth, and P. A. Miller,

- 2005: Architecture of MADIS data processing and distribution at FSL. Preprints, *21st Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, San Diego, CA, Amer. Meteor. Soc., CD-ROM, 2.39.
- Miller, P. A., M. F. Barth, and L. A. Benjamin, 2005: An update on MADIS observation ingest, integration, quality control, and distribution capabilities. Preprints, *21st Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, San Diego, CA, Amer. Meteor. Soc., CD-ROM, J7.12.
- Stensrud, D. J., 2006: NEH RTP Workshop: Improving weather forecast services used by the electric utility industry. *Bull. Amer. Meteor. Soc.*, **87**, 499–501.
- , and N. Yussouf, 2003: Short-range ensemble predictions of 2-m temperature and dewpoint temperature over New England. *Mon. Wea. Rev.*, **131**, 2510–2524.
- , and —, 2005: Bias-corrected short-range ensemble forecasts of near surface variables. *Meteor. Appl.*, **12**, 217–230.
- Yussouf, N., D. J. Stensrud, and S. Lakshmiarahan, 2004: Cluster analysis of multimodel ensemble data over New England. *Mon. Wea. Rev.*, **132**, 2452–2462.
- Zamora, R. J., E. G. Dutton, M. Trainer, S. A. McKeen, J. M. Wilczak, and Y.-T. Hou, 2005: The accuracy of solar irradiance calculations used in mesoscale numerical weather prediction, *Mon. Wea. Rev.*, **133**, 783–792.