

FORECASTING TECHNIQUES

**The Use of Hourly Model-Generated Soundings to Forecast Mesoscale Phenomena.
Part I: Initial Assessment in Forecasting Warm-Season Phenomena**

ROBERT E. HART AND GREGORY S. FORBES

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

RICHARD H. GRUMM

National Weather Service, State College, Pennsylvania

16 December 1997 and 14 July 1998

ABSTRACT

Since late 1995, NCEP has made available to forecasters hourly model guidance at selected sites in the form of vertical profiles of various forecast fields. These profiles provide forecasters with increased temporal resolution and greater vertical resolution than had been previously available. The hourly forecast profiles are provided for all of NCEP's short-range models: the Nested Grid Model, Eta Model, and Mesoscale Eta Model. The high-resolution forecasts aid in the timing of frontal passages, low-level jets, and convective initiation. In addition, through time–height cross sections of Richardson numbers, forecasters can alert pilots to the potential for clear air turbulence several hours to a day in advance. Further, the profiles are useful in prediction of cloudiness and the dissipation of low-level stratus and fog. Time–height cross sections of wind velocity have proven extraordinarily useful in visualizing and forecasting inversion heights, frontal passage timing, boundary layer depth, and available environmental and storm-relative helicity during convective events.

The hourly model forecasts were found to be exceptionally helpful when combined with hourly surface observations to produce enhanced real-time analyses of convective parameters for use in very short term forecasting. High-resolution analyses of lifted index, CAPE, convective inhibition, moisture flux convergence, and 2-h changes in these fields aid the forecaster in anticipating convective trends. The introduction of model forecast error into these real-time analyses was minimized by using the latest available Eta or Mesoscale Eta Model runs. Therefore, the model data used to enhance the analyses are typically no more than 6–12 h old.

1. Introduction

One of the greatest challenges facing forecasters today is the need to extract essential information from a great wealth of available data. Until recently, model guidance available to forecasters was generally on grids no finer than 100 km in resolution with 5–10 vertical layers and at intervals of 3–6 forecast hours. During the past five years, however, the volume of available gridded data has increased ten-fold and continued growth is anticipated. The Eta (Black et al. 1993) and Mesoscale Eta Models (MESO; Black 1994), the National Center for Environmental Prediction's (NCEP's) most complex synoptic model and first mesoscale model, respectively, produce high-resolution output at 30–50 vertical levels

and at every forecast hour in the form of soundings (or “profiles”). Therefore, the fine time and spatial resolutions of the model output make possible operational prediction of mesoscale features unlike what has been possible previously. The forecasting problem is how to use this high-resolution data to improve the forecasting process without overwhelming or distracting the forecaster.

The dataset used in this research is relatively new and few methods have been developed previously to visualize the data. Clearly none took advantage of the data to its fullest. The primary method for visualization of model soundings has been through BUFKIT, a package developed locally at the Buffalo office of the National Weather Service (NWS) (Niziol and Mahoney 1997). This exceptional package was developed primarily for lake-effect snow forecasting. BUFKIT aids forecasters through animations of hourly forecast soundings and maps of projected lake-effect snowbands based on the wind profiles and inversion heights within those soundings. The latest version of BUFKIT also provides fore-

Corresponding author address: Mr. Robert E. Hart, Department of Meteorology, The Pennsylvania State University, 503 Walker Bldg., University Park, PA 16802.
E-mail: hart@ems.psu.edu

TABLE 1. Database period and duration of model sounding profiles for each model available. The archive duration of the 0900 UTC run of the MESO is shorter than for other models because it was run solely for the support of the 1996 Summer Olympic Games. The NGM archive was locally halted in late 1996 because there was insufficient CPU time to process all three models in a real-time scenario. The NGM archive was resumed in September 1997 after a hardware upgrade.

Model	Run	Archive period	Archive duration (months)
NGM	0000, 1200 UTC	September 1995–December 1996	16
Eta	0000, 1200 UTC	February 1996–current	17
MESO	0300, 1500 UTC	March 1996–current	16
MESO	0900 UTC	May 1996–August 1996	3

casters with many of the graphical formats used in this research, including time–height cross sections and time series histograms.

For several years another software package, GEMPAK (desJardins and Petersen 1985), has been used to draw model-forecasted soundings based on the 3D grids. The user picks a location within the 2D domain of the model grid and GEMPAK plots the interpolated sounding profile based on the gridded model output. The advantage of this approach is that the user can choose any location within the region (the user is not limited to a fixed set of stations). The disadvantages are significant, however. The GEMPAK sounding typically has a vertical resolution of 25–50 mb, since the raw model grids used to generate the profile are only available at this spacing. Shallow layers containing potential significant information concerning convective instability, precipitation type, or jet streams may be missed due to this coarse vertical resolution. Second, the grids are produced every 3–6 h of the forecast period, limiting the temporal resolution of the interpolated soundings.

The purpose of this paper is to demonstrate how high-resolution model data can be used to improve the forecasting of mesoscale phenomena. The results show how these data displayed using visualization software can be used to improve forecasts of convection, turbulence, temperatures, and fog. This paper is divided into four sections. Section 2 is a detailed description of the methodology involved in this research, including the datasets, hardware, and software used. Section 3 examines the results of the research. A concluding discussion is given in section 4, which presents significant findings and avenues for future research. This article focuses on

the application of model soundings to forecasting warm-season phenomena. The second article in this series (Hart and Forbes 1999, manuscript submitted to *Wea. Forecasting*, hereafter HF) examines their utility in forecasting nonconvective strong wind gusts.

2. Methodology

a. Data

The data used were hourly forecast vertical profiles in Binary Universal Form for the Representation of meteorological data (BUFR) format. These BUFR profiles were retrieved from the anonymous file transfer protocol (FTP) servers at NCEP and the Office of Science Operations (OSO) and decoded locally into tabular format. Hourly profiles were retrieved for each of NCEP's short-range operational models: Regional Analysis and Forecast System Nested Grid Model (NGM; Hoke et al. 1990), Eta (Black et al. 1993), and MESO (Black 1994). The hourly profiles are full-model-resolution vertical profiles of model output, produced at every forecast hour. Data are available for temperature, wind velocity, precipitation, moisture, and parameterized or diagnosed values such as skin or 2-m temperature. Each of the model runs and corresponding data archive duration are illustrated in Table 1.

Unlike the NGM, the Eta and MESO are not fixed models. The Eta and MESO's physics and parameterizations are being improved with time. Several of these improvements came as a result of this research and through communication with scientists at NCEP. These model changes can have a dramatic impact upon the interpretation of these results. Table 2 summarizes the changes that occurred during this research, and the dates of their implementation.

NGM model soundings are obtained through interpolation between grid points to the exact observing station location. In contrast, the Eta and MESO forecast profiles are taken from gridpoint locations and are not interpolated to the exact surface observation station location. Consequently, there is a displacement between the hourly model profile location and the surface observation location for the Eta and MESO Models. Using the closest available grid point, the maximum displacement possible is approximately 34 km for the Eta and 20 km for the MESO. This misalignment could often

TABLE 2. Summary of relevant changes to the Eta and MESO Models that impact the results of this research.

Date	Description of model change
August 1995	NGM soundings made available
January 1996	Eta and MESO soundings made available
February 1997	Radiation parameterization scheme fixed (ozone and aerosol concentrations)
March 1997	Several dozen model sounding stations added to Eta and MESO
March 1997	Actual land-based model sounding grid points located over water are moved to have corresponding model grid points over land

be detected when precipitation totals associated with the model soundings were compared to those disseminated in the NCEP Forecast Output User Statistics (FOUS) data, where spatial interpolation within grids was performed. One consequence of the misalignment described above is that for land-based stations near a lake or ocean, the closest grid point to that station may be over water.

In early 1997, this mislocation problem was recognized by scientists at NCEP (Table 2). At that time, all stations suffering from the land–ocean/land–river mislocation problem were assigned to a different grid point. Thus, prior to this change stations such as Erie, Pennsylvania; Portland, Maine; and Key West, Florida, were associated with model soundings that were over water. After the change, the associated model soundings were shifted to grid points located over land. While this alleviates, to a considerable degree, the problems just discussed, it creates another: the mislocation distance between the station and model sounding has dramatically increased for these stations, some by a factor of 2–5.

The degree of station–model sounding mislocation is quantified in Fig. 1. For each model sounding, the associated surface observation station is marked with an “×.” The associated Eta and MESO Model sounding grid points are also shown. The surface elevation differences between the observation stations and the model soundings are depicted. One can note immediately that the model sounding grid point is not the same from model to model, the result of varying grid spacing and gridpoint location. In addition, certain sites had their model sounding grid point deliberately misplaced to match Next Generation Radar (NEXRAD) locations rather than the actual surface observing site. Cases of this include Cincinnati, Ohio; Wilmington, North Carolina; Dover, Delaware; and Morristown, Tennessee. Further, certain sites (WHI, C26) were chosen to verify wind profiler locations and thus do not have a corresponding NWS surface observing site. For this reason, these stations have not been included in Fig. 1. There are several forecasting consequences of the displacements, which will be discussed throughout the results section.

b. Methods

Once the forecast soundings are in ASCII tabular format, a series of forecast products are then generated using the Grid Analysis and Display System (GrADS; <http://grads.iges.org/grads>). Once the data has been retrieved from the FTP server at NCEP or OSO, the entire suite of forecast products (for approximately 25 sites) is generated in 2–3 h, permitting real-time use of the forecast products by forecasters at the National Weather Service Office in State College, Pennsylvania, and elsewhere. The forecast products, along with their applications, advantages, and disadvantages, are summarized in Table 3. This table summarizes applications of these

products that cannot be fully explained here but were observed frequently during the course of this research. Lastly, the hourly model sounding data are integrated with real-time surface data to produce hourly analyses of convective parameters such as CAPE, lifted index, and convective inhibition. The greatest hindrance to efficient product generation was server downtime and slowdown at NCEP, which could often delay product generation by 3–5 h or lead to a loss of current data once or twice a week.

While GrADS provided nearly all graphical requirements for the creation of forecast products, there was no skew T – $\log p$ capability in the GrADS package at the time the research was started. To overcome this limitation, a GrADS function was developed that allows users to create such diagrams. This is a fully functional skew T – $\log p$ routine including capability to plot wind profile, stability indices, hodograph, storm-relative helicity, and parcel traces. In addition, the function allows for sounding overlay so that model forecast values can be directly compared to observed soundings, Doppler radar–derived wind profiles, or other model output.

The development and rapid expansion of the World Wide Web during the early stages of this research provided an excellent interface by which nearly all users could acquire the graphical forecast products. In early 1996 a form-based Web interface was developed to provide users access to the forecast images. Through this Web site (<http://www.ems.psu.edu/wx/etats.html>), users can also monitor the status of the datasets to determine when new data has been received and also to browse any dataset received during the past 24 h.

3. Results

The forecast model sounding profiles for approximately 100 stations were archived for a period of time ranging between 3 and 17 months, depending on the model (Table 1). This extended archive allowed for detailed statistical analysis of the forecast soundings and their ability to predict various conditions. These results examine the utility of hourly model profiles in forecasting warm-season phenomena (late spring through early fall): shelter temperature, fog, turbulence, stratocumulus burnoff, and thunderstorm potential. Greater detail on all these results can be found in Hart (1997).

a. Two-meter temperature histograms

Prior to the operational implementation of the Eta Model, forecasts of shelter temperatures could be derived through statistical approaches only, such as the current Model Output Statistics (MOS) approach (Dallavalle et al. 1992). In such an approach, statistical forecasts of shelter temperature were developed based on forecast temperatures and thicknesses at various levels within the atmosphere. The advantage of such an approach was that it accounted for model biases effec-

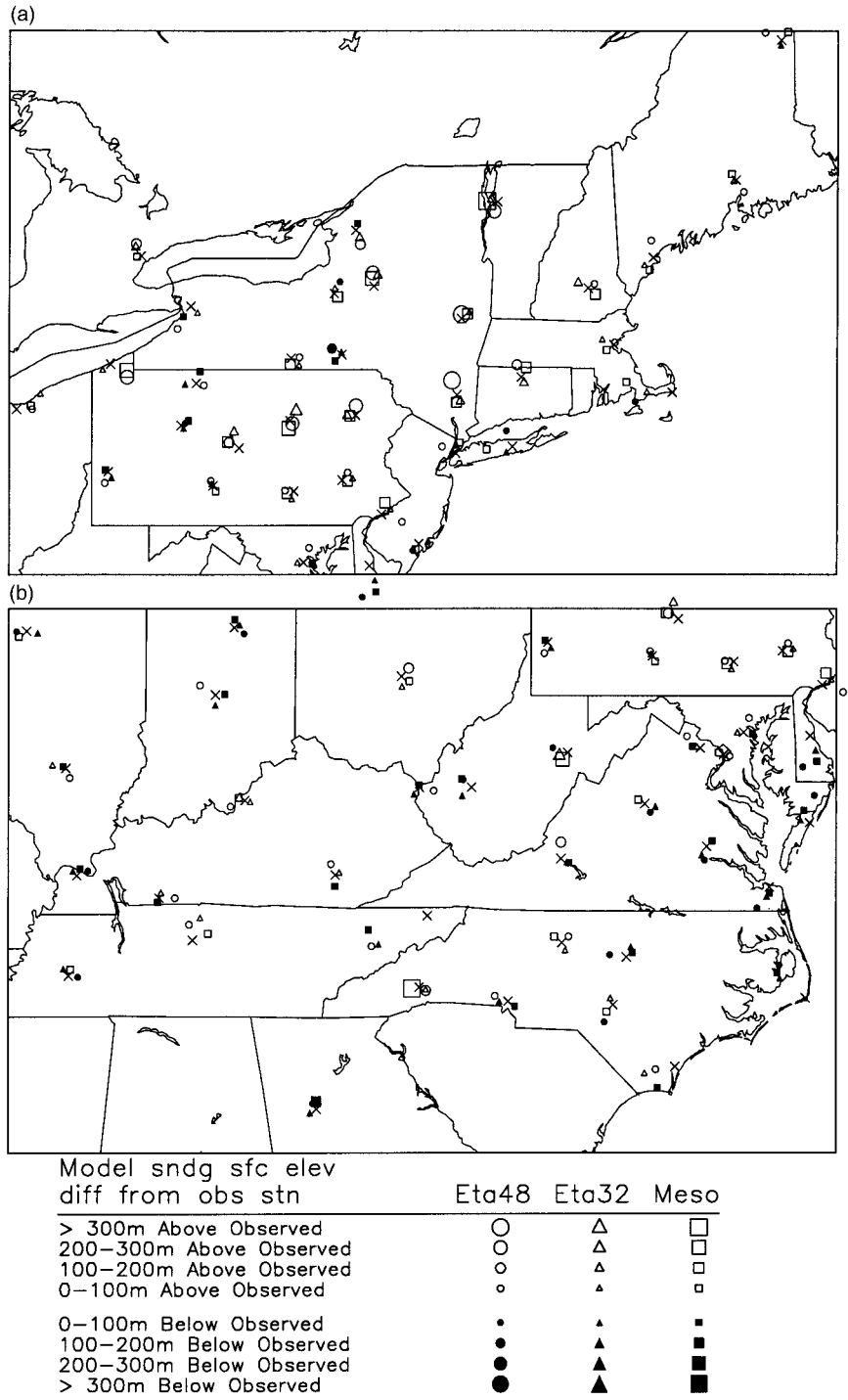


FIG. 1. Surface station model sounding gridpoint mislocation magnitude for several stations. The crosses represent the surface observation station location. The legend indicates the symbols used to measure the degree of vertical mislocation between the model sounding and the observing station. This mislocation (both horizontal and vertical) is significant to the forecast since it represents a consistent bias that exists between what is numerically forecast and what is observed. The bias is reflected in almost all operational and experimental output, including precipitation type, wind gust forecast, surface temperature forecast, and convective potential. Further, the mislocation varies between the models, and therefore when forecasters perform model comparison, this bias must be subjectively accounted for.

TABLE 3. Summary of types of forecast products generated from the hourly model soundings. Forecast applications of each of the products as well as each product's advantages and disadvantages are presented.

Plot type	Profile graphical display products			
	Display fields	Forecast applications	Advantages	Disadvantages
Time–height series	Temperature Relative humidity Vertical velocity Equivalent potential temperature Wind profiles Richardson number	High and low temperature Cloud type and level Wind speed and direction Turbulence	Temporal changes for a single location during forecast period Trends and progression of sensible weather at a single location	Subject to model bias Subject to model timing errors No sense of spatial forecast
Skew T –log p plots	Static soundings Animated soundings	Stability (changes) Clouds Winds Precipitation and type	Point in time for a specific location Detailed changes in temperature and moisture	Subject to model bias Subject to model timing errors No sense of spatial forecast
Distance–height cross sections	Static and animated Temperature and meridional winds Relative humidity Vertical velocity Potential temperature	Cold air damming Frontal timing Timing precipitation Clouds	Interpret for points between forecast locations Visualize weather features	Subject to model bias Subject to model timing errors
Time histograms	Accumulated precipitation and type 2-m temperature, dew-point, and skin temperature Fog product Convective parameters lifted index K index total totals Storm-relative helicity Frost and freeze products Wind gust probability	Precipitation Temperatures Convection severity type timing location Agricultural frost warnings freeze warnings High wind watches–warnings	Refined local forecasts for specified points Specialized products provide guidance on potential weather elements Easy to interpret and user friendly	Subject to model bias Subject to model physics and parameterization changes Subject to model timing errors No sense of spatial forecast

tively. The disadvantage was that one equation was used per forecast variable during a whole season, thus inately limiting the usefulness of the approach during anomalous situations.

The Eta and MESO Models have more complex parameterizations than the NGM, and the output provided to the user by these two models has increased. The model outputs forecast 2-m temperature fields that, for practical purposes, serve as shelter temperature forecasts. Since parameterized physical processes determine the 2-m temperature, the 2-m temperature forecasts should be superior to the MOS-based shelter temperature forecasts during anomalous situations. Indeed, we found that during severe arctic outbreaks and heat waves, the 2-m temperature more accurately forecasted observed shelter temperatures than MOS. The most notable increase in accuracy was seen during sharp temporal changes in temperature, which the MOS approach tends to smooth out with time.

There are significant problems with using the 2-m temperature as a shelter temperature forecast. Statistics were compiled for 12 model sounding stations in the northeast United States on the use of 2-m temperature as a shelter temperature forecast for a 1-yr period (March 1996–February 1997). These statistics were produced for both the Eta and MESO Models. The NGM

2-m temperature forecast is not provided. Figures 2 and 3 present the bias and root-mean-square error (rmse) for each of the 12 stations for both models. Figure 2 plots the statistics as a function of forecast hour, while Fig. 3 plots the statistics as a function of time of day.

For 10 of the 12 stations there was a strong cold bias during the nighttime and a weak to strong warm bias during the daytime. The cold bias peaked just before sunrise and the warm bias peaked during maximum heating. The magnitude of the bias at rural locations (Altoona, Pennsylvania; State College, Pennsylvania; and Binghamton, New York) was less than in urban areas. This suggests that the effects of heat-island influence were contributors. Clearly, however, this cannot be the sole cause since the bias still exists in rural areas. The bias indicates that the surface radiation and/or boundary layer parameterization schemes may be in error. Another source of the bias would be elevation mislocation between the sounding and station, as previously mentioned. The largest forecast errors occurred during times of minimal cloud cover. When a clear night was forecasted, the model 2-m temperature forecast was, on average, 3°–6°C too low. The forecast error was less during times of cloud cover and precipitation.

In February 1997 (Table 2), NCEP scientists discovered a problem with the radiation schemes within the

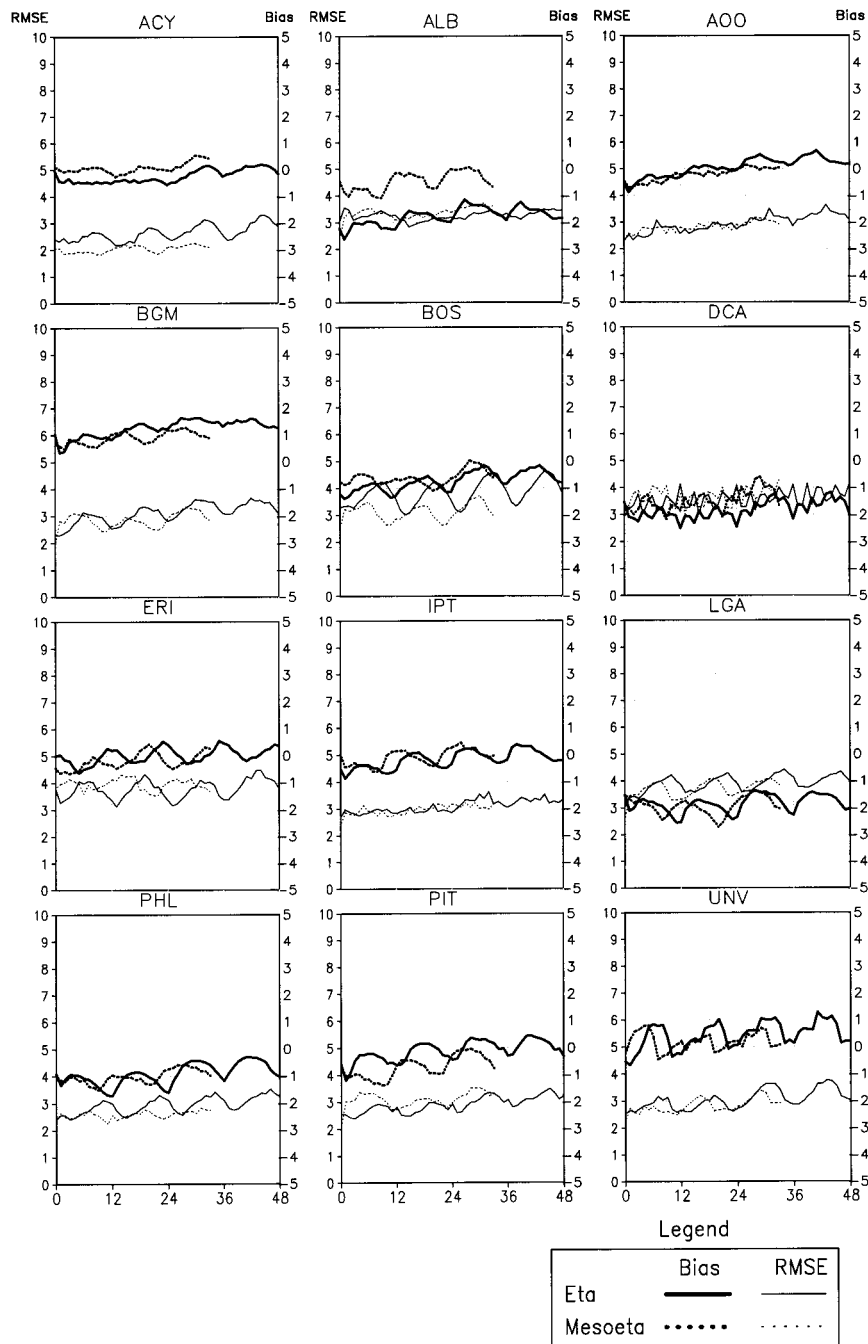


FIG. 2. Eta and MESO 2-m temperature bias and rmse ($^{\circ}\text{C}$) as a function of forecast hour for 12 Northeast stations.

Eta and MESO Models. The radiation schemes were producing excess net shortwave radiation at the ground. This excess was the result of three problems: the model used a noncentric orbit of the earth around the sun, an erroneous ozone parameterization, and erroneous aerosol parameterizations. In addition to the biased daytime shelter temperatures that result from the excess shortwave radiation, the radiation schemes were pro-

ducing more vigorous boundary layer mixing that generated too deep a mixed layer, lowering the shelter dew-points. These parameterization errors were producing the 2-m temperature biases described earlier. The cause of the 2-m temperature forecast error also explains why the bias was greatly reduced during cloudy days.

One station, Erie, Pennsylvania, had the inverse relationship with a cold daytime bias and a warm night-

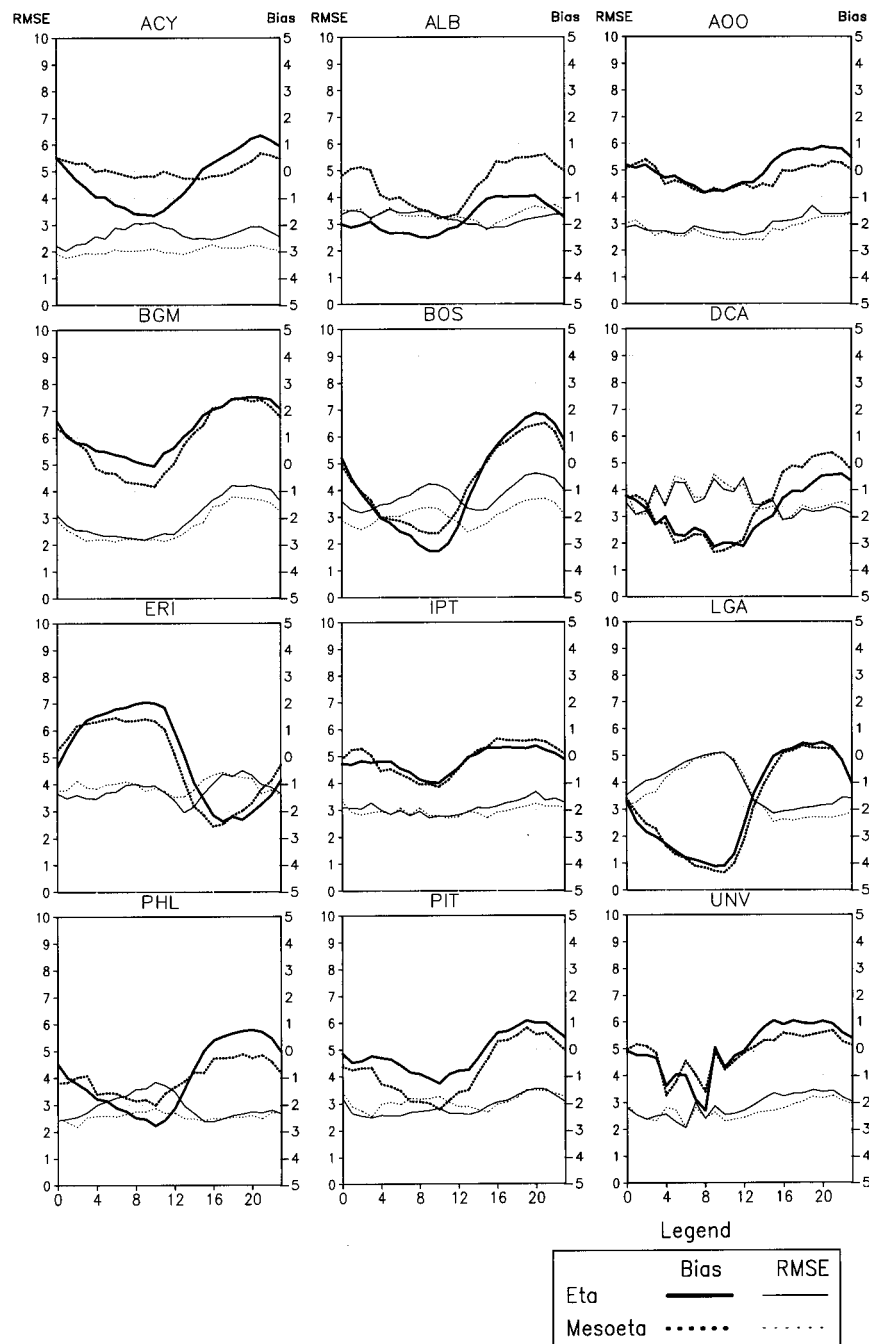


FIG. 3. Eta and MESO 2-m temperature bias and rmse ($^{\circ}\text{C}$) as a function of time of day (UTC) for 12 Northeast stations.

time bias (Fig. 3). The cause of this is the land–ocean misplacement described previously. The forecast sounding for Erie was located over water for both the Eta and MESO Models. Once again, however, it should be noted that all stations now have corresponding model sounding stations that are located over land and thus the bias presented in Erie and other coastal stations should be diminished. Since these parameterizations have been

fixed, it has been observed that the shelter temperatures forecast in the model soundings have improved considerably. Preliminary poststudy analyses show that the 2-m temperature forecast by the Eta and MESO Models now have a slight to moderate (1° – 4°C) cool daytime bias and a slight (1° – 2°C) nighttime warm bias.

While a discussion of the full ramifications of this low-level temperature bias is beyond the scope of this

TABLE 4. Optimal threshold low-level relative humidity values for prediction of fog (visibility less than 2 mi) for each of the three models. The forecast accuracy at each of these threshold values is also presented.

Model	Optimal threshold relative humidity for fog prediction (visibility \leq 2 mi)	Forecast accuracy (%)
NGM	95	66
Eta	92	61
MESO	93	65

paper, it is interesting to note that these biases may contribute to an increased coastal area baroclinicity. For example, in the polar or arctic air mass over land, low-level temperatures are likely to be too cold in the Eta and MESO, especially during the night, as discussed earlier. Over the water, especially the Gulf Stream, this bias does not exist since the skin temperature is held fixed to the temperature of the water, thus preventing excessive radiational cooling as is seen over land during clear nights with these models. Even over land in the presence of a tropical air mass, the low-level parameterized temperature bias is likely to be greatly diminished (compared to that in dry arctic air masses) since the moisture content of the boundary layer is excessive. Consequently, these models may artificially enhance the low-level baroclinicity along the East Coast as a result of this gradient of low-level temperature bias. This increased coastal baroclinicity may lead to an anomalous enhancement of the deepening rates during cyclogenesis.

b. Experimental fog forecast

The ability of a model to predict fog is directly related to its ability to predict low-level moisture. Not all models predict variables explicitly at the surface; therefore, a level close to the surface must be chosen to represent the conditions at the surface. While fog usually occurs at relative humidity of close to 100%, model parameterizations of explicit and convective precipitation generation work feverishly to remove saturation. As a result, it is necessary to determine what forecast subsaturation low-level relative humidity can be associated with a prediction of fog. With this question in mind, hourly model sounding data were analyzed to determine the ability of the profiles to predict fog. The results of this analyses are presented in Table 4.

Based on analyses of several hundred forecast hours, it was found that an NGM forecast relative humidity of 95% or higher in the lowest sigma layer suggested a forecast of visibility of 2 mi or less. Data were pooled for 11 cases of synoptically forced fog and low visibility at Pittsburgh and State College, Pennsylvania, associated with periods of precipitation from September 1995 through March 1996. On average, the forecast was correct 66% of the time, on an hourly basis, for all forecasts

out through 48 h. Surprisingly, NGM forecasts valid from 0 to 12 h in advance were somewhat less accurate than average, and forecasts for hours 30–42 were somewhat more accurate than average. It was apparent that in many instances the hourly changes in NGM relative humidity gave indications of the trends toward increasing or decreasing visibility. Further, the NGM forecasts did not always verify observed fog observations on an hour-by-hour basis, but they did show considerable skill in indicating that fog was likely to develop, persist, or dissipate during a period of a few hours through hourly changes in low-level relative humidity. The NGM hourly soundings did not appear to possess significant skill in forecasting localized ground fogs due primarily to radiation.

MESO sounding output for Altoona, Pennsylvania, and Pittsburgh, Pennsylvania, were examined during May 1996. Upon examination of three fog events (8 days), it was found that the 2-m relative humidity is a better predictor of fog than is the relative humidity in the lowest model layer. The 2-m relative humidity was calculated using the 2-m temperature and 2-m mixing ratio. It was determined that the appropriate MESO threshold forecast relative humidity for a prediction of fog of visibility 2 mi or less is 92%. When used, this forecast was correct 61% of the time, compared to 54% when a relative humidity threshold of 95% was used. The trend of the 2-m relative humidity was often an excellent indicator of the anticipated breakup or formation of fog. Forecast errors were more often due to timing errors than otherwise.

Another indicator of fog breakup or formation was often found by comparing the 2-m relative humidity to the lowest model layer relative humidity. In several cases, fog formation was associated with a trend in which the 2-m relative humidity was increasing faster than the lowest model layer humidity. Conversely, fog breakup was often associated with a trend in which the 2-m relative humidity was decreasing more rapidly than the relative humidity of the model layer above it. There were not sufficient cases of these events to quantify the forecasting ability of this trend.

A similar procedure was applied to the Eta soundings between February and May 1996. Once again, the 2-m relative humidity was found to be a more accurate indicator of fog than the lowest model layer above it. The optimal threshold for fog prediction (again, visibility 2 mi or less) by this model was found to be 93% (at the 2-m level). Forecasts using this threshold were correct 65% of the time. Similar to the MESO model, the trend of the fog was often indicated by the trend in the relative humidity at both the 2-m and lowest model layers. As with the NGM soundings, neither the MESO soundings nor the Eta soundings appeared to have success in predicting ground fogs.

c. Experimental turbulence forecast histograms

The accuracy of Richardson number forecasts was examined for the Eta and MESO Models for a period

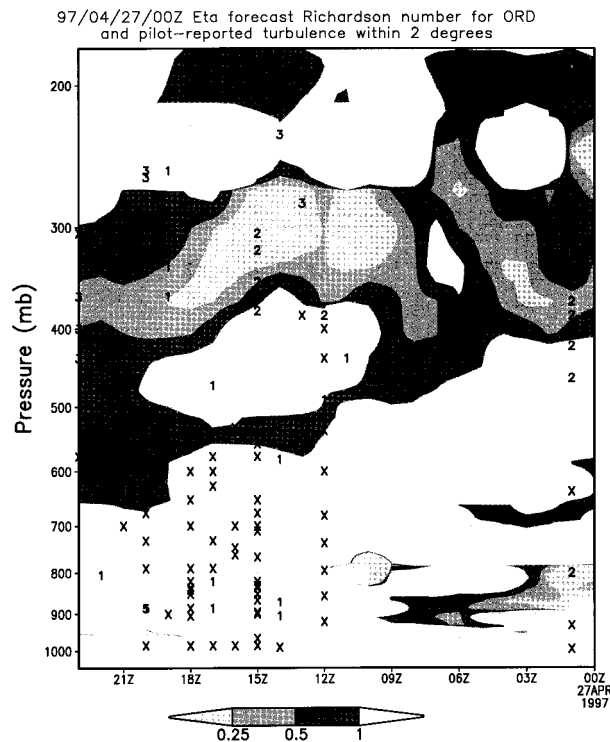


FIG. 4. Example overlay of forecast Richardson number and pilot observed turbulence. Analysis of this field was used to evaluate the ability of the Eta and MESO Models to forecast CAT. The shaded field is the forecast Richardson number and the superimposed numbers are the magnitude of pilot-reported turbulence. An “x” indicates that a pilot did not report turbulence.

of 2 weeks (3–16 April 1997) at three major airports (Chicago’s O’Hare, New York City’s LaGuardia, and Pittsburgh). This period of time is advantageous in that it minimizes potential data contamination from convectively produced turbulence. Pilot reports (PIREPS) were examined and these reports of turbulence were used to verify Richardson number forecasts from the first 24 h of each model forecast. For each station, a pilot report was deemed “at the airport” if it occurred within 2° lat–long of that airport. Pilot reports of turbulence (with intensity estimates rated 1 to 7, for weak to severe, respectively) were superimposed on model forecast fields of Richardson number, as shown in Fig. 4. If a pilot did not report turbulence, an “x” was marked on the diagram to attempt to validate regions predicting the lack of turbulence. From a model forecast standpoint, turbulence was forecasted as possible when the Richardson number dropped below 1 and probable when it dropped below 0.25. The results of this analysis are shown in Table 5.

Pilot report density played a significant role in verification of turbulence forecasts. Flight occurrences drop off dramatically after 0100 UTC and remain minimal through 1100 UTC. During this time, it was difficult to determine whether an unverified forecast was the result of an inaccurate forecast or simply the lack of pilot

TABLE 5. Results from the analysis of Eta and MESO forecasting ability for pilot-reported turbulence. The period of analysis spans 2 weeks at three major airports—New York City’s LaGuardia (LGA), Chicago’s O’Hare (ORD), and Pittsburgh (PIT)—and only included forecasts of less than 24 h. During this period, 614 turbulent reports were sent by pilots. In both diagrams, the percentages represent the fraction of the 614 turbulent reports that occurred within a certain window of a forecast Richardson number. This window was defined as 2° horizontally, 100 mb vertically, and 3 h temporally. In (a), the percentages shown are for four independent ranges of Richardson number. (b) Successive cumulative totals of the bins in (a) are represented. For both models, nearly two-thirds of the time a pilot reported turbulence the forecast Richardson number was less than 0.25. Nearly 90% of the time turbulence was observed, the forecast Richardson number was less than 1.0.

Forecast Richardson number	Eta forecast vs pilot-reported turbulence (%)	Mesoscale Eta forecast vs pilot-reported turbulence (%)
(a)		
>1.0	14	11
0.50–1.0	13	15
0.25–0.50	7	10
<0.25	66	64
(b)		
<1.0	83	89
<0.50	73	74
<0.25	66	64

reports to verify it. As a result, forecasted turbulence occurring during this period was not examined unless a pilot report was available. Given the large verification radius for each airport (2°), a large verification window was applied for each airport. If a pilot report of turbulence occurred within 3 h or 100 mb of a forecasted event, then the report was considered valid for the event. In addition, the 100-mb vertical window was applied since Kelvin–Helmholtz wave generation as a result of clear air turbulence (CAT) may extend above the unstable layer (as a result of wave breaking or otherwise).

During the 2-week evaluation period, 614 pilot reports of turbulence were observed. Of those reports, 66% were associated with Eta-based Richardson number forecasts of less than the empirical critical threshold of 0.25. If the constraint on verification is lessened, 73% of the reports occurred with a forecast Richardson number of less than 0.5 and 86% occur with a forecast of 1 or less. The ability to forecast turbulence does exist, with nearly two-thirds or more of observed turbulence events occurring within the empirically calculated threshold.

Determining the false alarm rate of turbulence forecasts is difficult since the density of pilot reports is low. The vast majority of turbulence reports occur within 200 mb of the surface, with a fraction occurring above this layer and then mostly at cruising altitudes near jet streams. The number of pilot reports in the middle troposphere (700–400 mb) is low, giving low confidence to forecasting ability in this region of the atmosphere. Subjectively, it can be said that most of the time when a potentially turbulent layer (Richardson number < 1)

was forecasted above the boundary layer during the 24-h period, a pilot report of turbulence was received. However, using the current observational dataset, the false alarm rate for predicting turbulence would be exceptionally high if the lack of a PIREP was used as a negative confirmation since the probability of a thin, short-lived turbulent layer being intercepted by a plane was low. For this reason, false alarm rate statistics are not being presented.

It was observed that model forecasts of boundary layer Richardson number often approach zero (and occasionally dropped slightly below zero during times of superadiabatic lapse rates) and yet turbulence was not reported if the turbulent layer was too shallow. From a verification standpoint, therefore, it is not practical to associate every critical layer with turbulence since extremely high false alarm rates would result. It is quite evident from this analysis that a minimal potentially turbulent layer depth is required in boundary layer forecasts (and perhaps aloft as well) for turbulence to verify, despite the duration of the unstable layer. For the three airports evaluated, this critical depth is approximately 75 mb. If this depth was exceeded, pilot reports of turbulence were almost guaranteed if the Richardson number dropped below 0.25. Further, if the potentially turbulent layer depth exceeded 150 mb, turbulence was always reported and often exceeded the rate of five reports per hour. Additionally, if the turbulent layer depth exceeded 150 mb, the pilot-reported turbulence would frequently reach moderate to severe levels (PIREP turbulence levels 3–7). Yet, if the layer depth dropped near or below 75 mb, the number of pilot-reported turbulence events would dramatically decrease or vanish. If the depth dropped to or below 50 mb, it was exceedingly rare for a pilot to observe turbulence (even weak turbulence) near or in the boundary layer, even if the Richardson number was near zero for the entire depth of the shallow layer. Given the large number of pilot reports in and just above the boundary layer, there is considerable confidence in these statistics involving critical depth.

Statistics and results from the MESO forecasts are similar to those from the Eta forecasts. For the 614 pilot-reported turbulence events, 64% occurred with a Richardson number below 0.25. Further, 74% of the events occurred with a Richardson number of less than 0.5, with 89% occurring with a Richardson number of less than 1, as shown in Table 5b. The slightly increased accuracy may be due to the increased vertical resolution of this model; this enables stronger vertical gradients of wind and temperature, allowing for increased likelihood that pilot-reported turbulence would verify in the MESO than in the Eta Model, where such gradients are slightly less resolved. Regardless, it appears from this examination that further increases in model vertical resolution will not dramatically increase the forecast accuracy of turbulence, although it may make the forecasts more precise. Doing so, however, also raises the potential for false alarms, since the areal (vertical and temporal) cov-

erage of the 1.0 value of Richardson number is likely to increase as vertical resolution increases.

d. Low-level moisture, stratus, and stratocumulus burnoff

During the warm season, a very significant forecast problem became evident for Pennsylvania and the surrounding region: the Eta and MESO Models have extreme difficulty in reliably predicting removal of early morning low-level stratus and fog. It was quite common that the model would predict *rapid burnoff* of the moisture (boundary layer relative humidity dropping below 80%), and yet it never occurred, or occurred near sunset. This has severe implications on forecasts of insolation, shelter temperatures, and instability. If the 2-m temperature is used as a forecast of shelter temperature, the largest daytime forecast errors occur between 1700 and 2100 UTC (Fig. 3). Of the 10 days with largest forecast 2-m temperature error during the spring and summer of 1996, 9 of the 10 had overcast skies and fog as weather conditions. On each of these days, the forecast 2-m temperature was off by 10°–20°C. Such forecast errors wreak havoc on forecasters' ability to predict the location and intensity of convection, especially when the environment is weakly forced. Is it possible to anticipate days when stratus and fog will not burn off as quickly as the model suggests by using forecast boundary layer temperature and moisture profiles?

In an attempt to answer this question, several dozen cases were examined during the spring of 1996 in Pennsylvania where fog and low-level stratus were reluctant to mix out while the Eta and MESO forecasted quick burnoff of the low-level moisture. It appears that in a majority of these cases, this phenomenon can be predicted with reasonable success by examining the vertical and temporal gradients of relative humidity within the boundary layer between 1200 and 1800 UTC. These gradients gave clear indications of the degree of mixing that was likely to occur and whether this mixing would be sufficient to remove the low-level moisture.

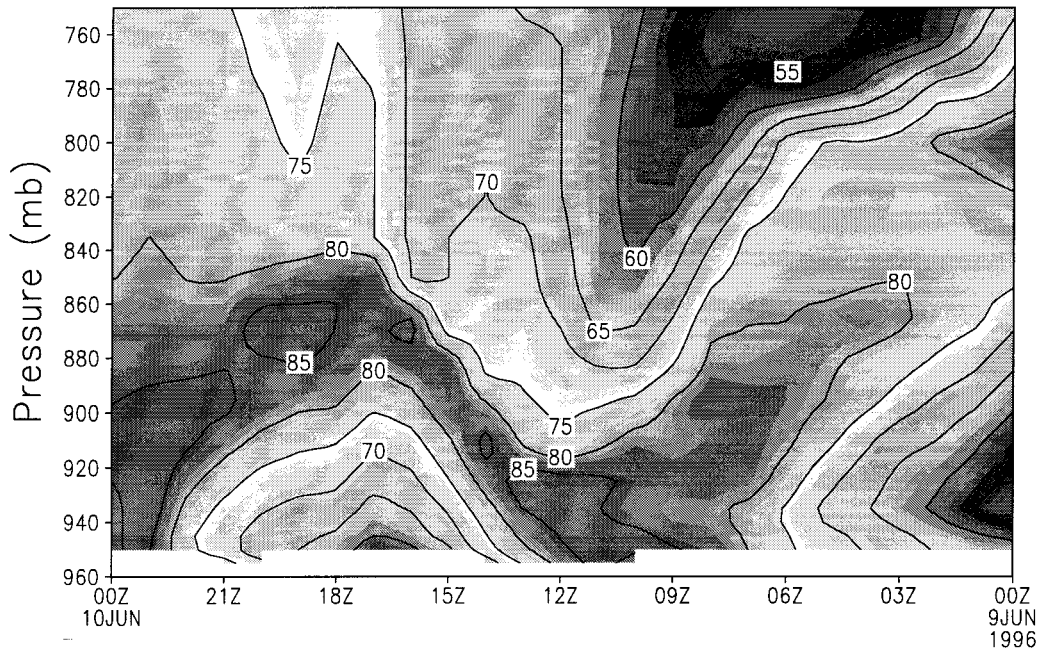
Results indicate that on 70% of the days, existing low-level stratus and fog will *not* break by 1800 UTC if the following two conditions are met prior to 1800 UTC.

- 1) The forecast maximum relative humidity within the boundary layer (lowest 100 mb) does not drop below 80% for two consecutive hours.
- 2) The forecast shelter (2-m) relative humidity does not drop below 60% for two consecutive hours.

Two examples of this approach are demonstrated in Figs. 5 and 6. The first example, Fig. 5, demonstrates the use of the approach to anticipate clearing during the morning hours. Figure 5a is a time–height cross section of forecast relative humidity from the 1500 UTC MESO run of 8 June 1996 for Altoona, Pennsylvania. The period of interest is from 1200 through 1800 UTC on 9 June.

(a)

1500UTC/June 8, 1996 Mesoeta Forecast RH for A00



(b)

Forecast 2-Meter and Observed Surface Temperature

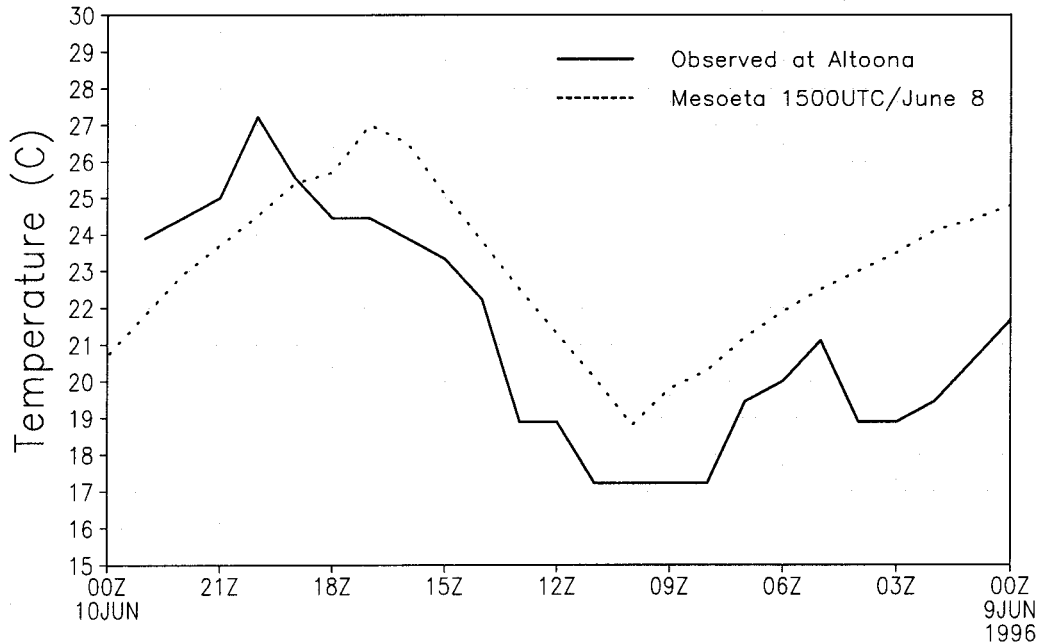
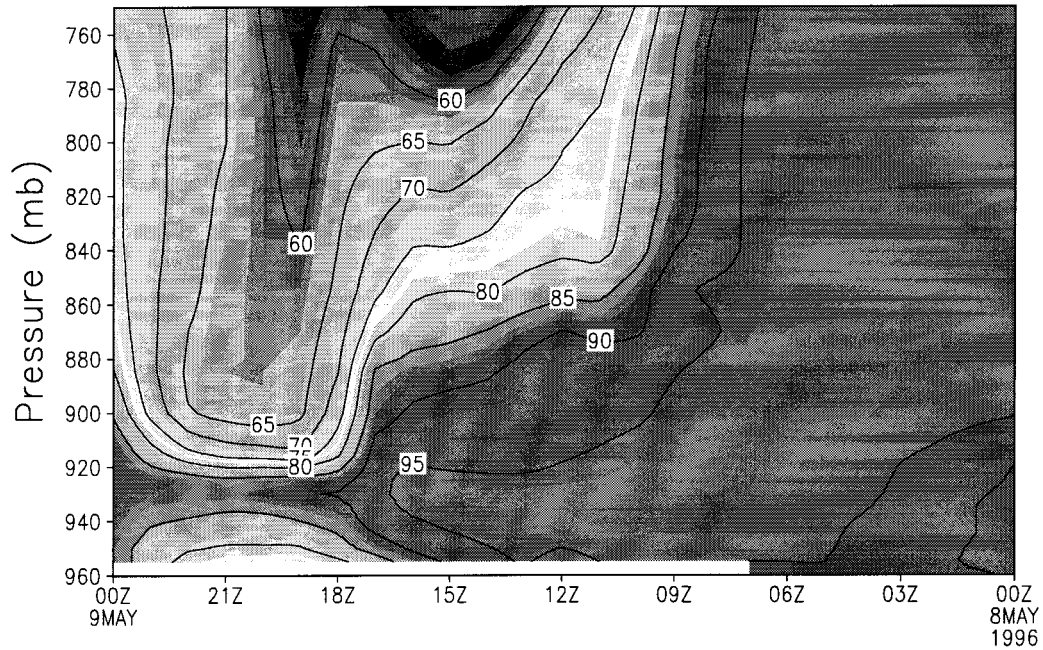


FIG. 5. Example application of (a) forecast relative humidity time-height cross section to anticipate model forecast error in (b) low-level temperatures. In this case, forecast 2-m temperatures, shown in (b), had considerable skill in predicting shelter temperatures during the afternoon of 9 June. The criteria specified in the text are not met in the time-height cross section of forecast relative humidity shown in (a). Thus, we conclude that sufficient mixing will indeed occur to allow low-level drying and allow forecasters to anticipate that the forecast 2-m temperatures will be realized.

(a)

1500UTC/May 7, 1996 Mesoeta Forecast RH for AOO



(b)

Forecast 2-Meter and Observed Surface Temperature

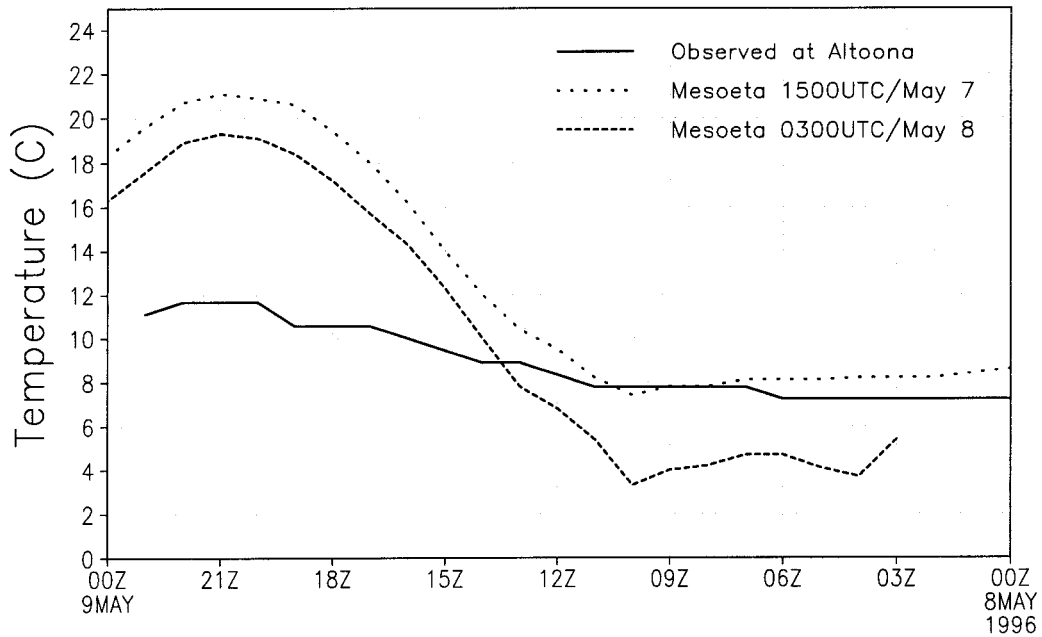


FIG. 6. Example application of forecast relative (a) humidity time-height cross section to anticipate model forecast error in (b) low-level temperatures. In (a), a forecast cross section of relative humidity is presented from the Mesoscale Eta Model for Altoona, PA. The model forecasts deep low-level moisture during the early morning of 8 May, which decreases greatly by early afternoon. In (b), the model forecasts 2-m temperatures to rise near 20°C by early afternoon. In reality, the observed shelter temperatures remained below 12°C the entire day and skies never cleared. By examining the temporal and spatial changes in moisture in (a) as described in the text, such a forecast error can be anticipated nearly 70% of the time.

At 1200 UTC the model forecasts a moist boundary layer, which then is mixed vertically through solar heating and the boundary layer dries to 55% relative humidity by 1700 UTC. However, as mentioned earlier, the model is often too quick in forecasting lowering of boundary layer relative humidity. Using the approach described above, it can be determined with reasonable accuracy whether the low-level relative humidity will, in fact, be decreased as the model indicates. The first criterion is clearly not met, since the relative humidity in the boundary layer between 1400 and 1800 UTC drops well below 80%. The second criterion is also not met since the shelter relative humidity drops below 60% at 1600 UTC. Forecasters should conclude that the model forecast for low-level stratus removal is accurate. The accuracy of the forecast on this occasion can be seen in the shelter temperatures. The 2-m forecast temperature (Fig. 5b) was a reliable forecast for the shelter temperatures that day, correctly predicting the maximum afternoon temperature, although the timing was off by 3 h.

The second example demonstrates the successful use of the approach to anticipate the maintenance of high boundary layer relative humidity and corresponding cooler shelter temperatures, avoiding a forecast bust. Figure 6a is the forecast relative humidity for Altoona, Pennsylvania, from the 1500 UTC 7 May 1996 MESO run. The time period of interest is 1200–1800 UTC on 8 May. A saturated boundary layer at 1200 UTC is forecast to remain moist, although below saturation, through 1800 UTC. Despite this, the forecast 2-m temperature (Fig. 6b) shows significant warming from 1200 through 1800 UTC during that day. Forecasters must be skeptical of such a high temperature forecast by the model, since the above two criteria are met for this period. Forecasters should expect that the low-level relative humidity will remain high, fog and stratus will not dissipate, and that shelter temperatures are likely to be considerably cooler than predicted. Indeed, as shown in Fig. 6b, the observed shelter temperatures rose only 4°C during the 6-h period, compared to the forecast of a 10°–11°C rise.

Examination of the 10 events similar to those just described show that on days when the fog and stratus do not dissipate by 1800 UTC during the spring–early summer, a typical daytime maximum temperature is about 3°–8°C warmer than the morning minimum. When stratus and fog dissipate early, a daytime warmup of 9°–20°C is typical. Thus, the approach just described can help forecasters reduce the forecast error of shelter temperatures during such events.

e. Thunderstorm potential forecasts

For each station available on the Web site, time series histograms of storm-relative helicity are provided to the forecaster, along with similar histograms of forecast lifted index, total totals index, and K index. Further, all of these variables, including convective inhibition, are made available on the sounding animations. Using these

products, the forecaster can quickly determine the degree to which forecast ingredients are coming together to yield severe weather at various stations in the forecast area. The ability of the Eta and MESO to forecast convective available potential energy (CAPE), convective inhibition (CIN), and helicity—three of the most crucial factors that determine the thunderstorm potential, type, and severity (Gaza and Bosart 1985; Johns and Doswell 1992)—are now examined. The forecasting ability of the Eta and MESO was analyzed during 18 convective events in the Pennsylvania and surrounding region during June and July 1996. During each of these events, the observed soundings at 1200 and 0000 UTC were compared to forecast model soundings for those times, using the three upper-air sites closest to the Pennsylvania region: Pittsburgh, Albany, and Washington D.C.'s Dulles International Airport.

1) CAPE FORECASTS

For each of the 18 convective events, the observed and forecast values of CAPE were compared. CAPE was determined based on the highest equivalent potential temperature layer in the lowest 250 mb at the three sites, which was typically the surface. These observed values were then compared to model forecast values of CAPE for the Eta and MESO Models. Results are shown in Fig. 7. The model forecast values for CAPE were those from the model run prior to the observation time, giving a 12-h forecast for the Eta Model and a 9-h forecast for the MESO Model. It is acknowledged that this analysis is not completely representative of the convective potential during these events since the observation times used (0000 and 1200 UTC) are typically numerous hours before and after the events occur and, thus, not representative of the convective environment during the most unstable time of the day in the United States. Regardless, this analysis gives a first glance at the ability of the Eta and MESO Models to predict these parameters, illuminating any biases that may be present and applicable to other times during the day when sounding observations are not taken.

In all four diagrams there is considerable spread in the data, indicating significant difficulty forecasting CAPE. At 0000 UTC, there appears to be slightly greater tendency for each model to overpredict CAPE than to underpredict it. In those cases where the models overforecast CAPE, it was typically the result of two factors: (a) forecasting the shelter temperatures too high as a result of erroneous cloud-cover forecasts described previously or the erroneous radiation parameterization described previously, and (b) the lack of cool downdrafts in the Betts–Miller convective parameterization scheme, (Black et al. 1993; Black 1994; Janjic 1994), which produces low-level temperature fields that are too warm following convection. The tendency to overpredict CAPE was limited by the models' tendency to underforecast afternoon dewpoints by 1°–2°C during these

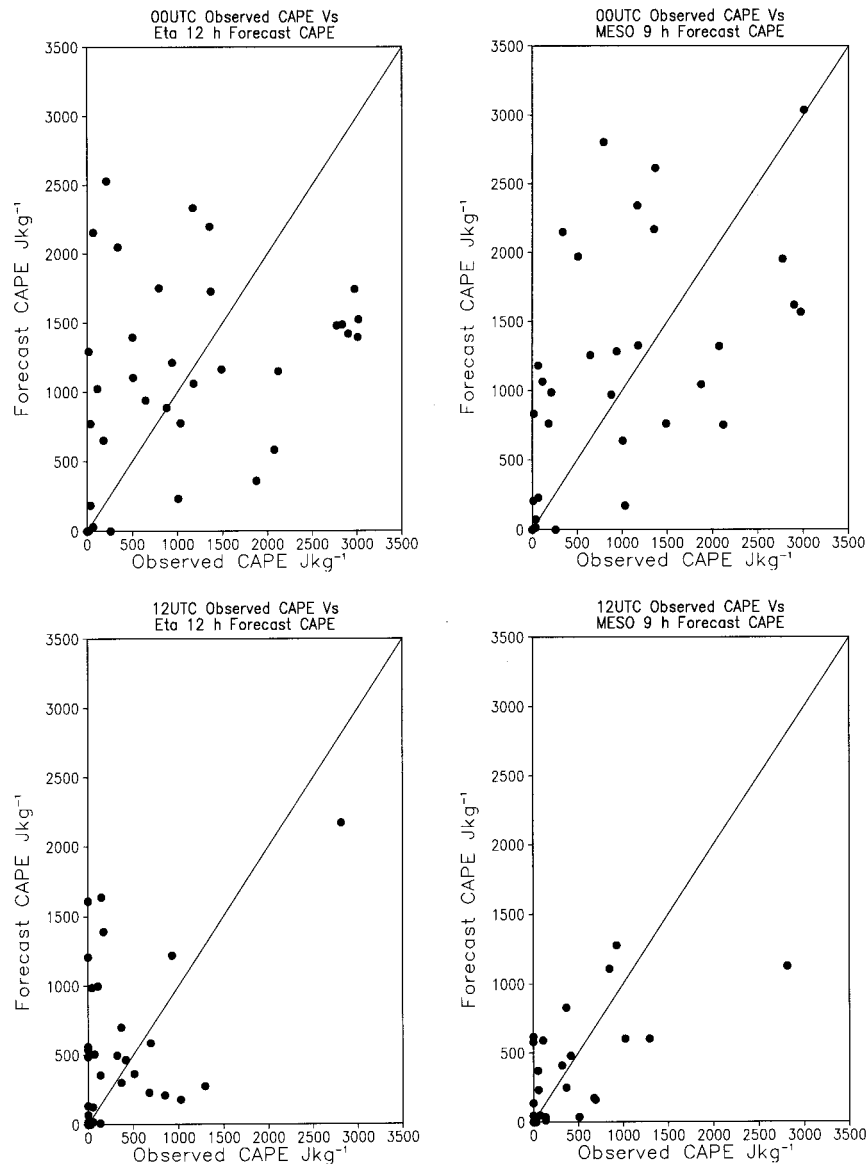


FIG. 7. Observed vs forecast CAPE for the Eta and MESO Models.

events. This dewpoint bias is consistent with the bias of overdeepening the boundary layer discussed previously. While the dewpoint has a greater impact on the CAPE calculation, the cool bias associated with the forecast dewpoints was considerably less than the warm bias associated with the forecast temperatures.

It was also noted on several occasions that when elevated convection was occurring in reality, the Eta and MESO Models had a tendency to underforecast the convective response and place the regions of convective rainfall too far into the low-level high equivalent-potential-temperature air. This results from the choice of convective parameterization in these models, since the scheme determines cloud base and vertical velocity based on the lowest 100 mb of the atmosphere, which

can often be quite stable or even exhibit an inversion during times of elevated convection. In such situations, the scheme would be more likely to produce precipitation in the low-level warm air, since the scheme responds more strongly if the unstable air is closer to the surface. Further discussion and examples of this convective bias are given in Grumm and Hart (1998, manuscript submitted to *Wea. Forecasting*).

At 1200 UTC, the MESO has considerably more skill than the Eta in predicting CAPE. During the 1200 UTC observations associated with these 18 events, the largest forecast error for CAPE was 1600 J kg^{-1} by the Eta but only 600 J kg^{-1} by the MESO. This can be partially attributed to the 3-h difference in forecast time between the two models. During several events, it was observed

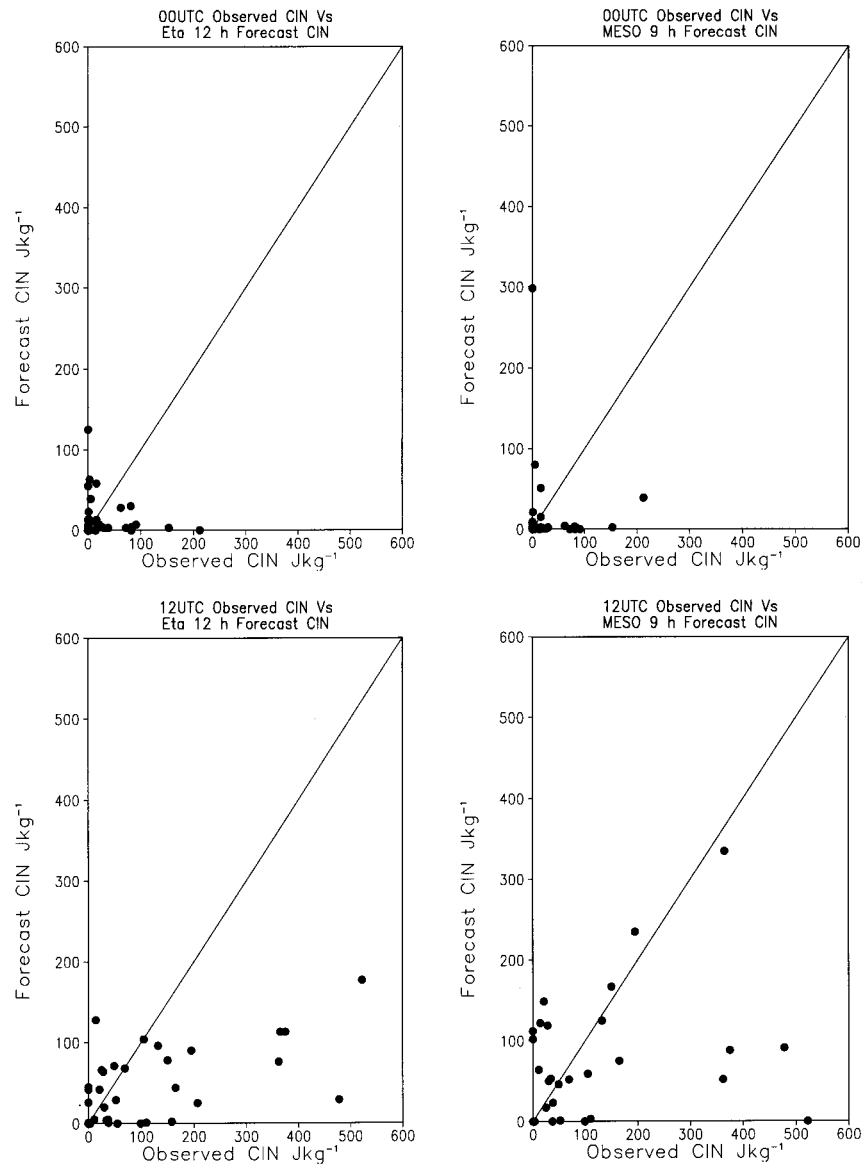


FIG. 8. Observed vs forecast CIN for the Eta and MESO Models.

that timing errors in frontal passage or advection patterns aloft were the cause of erroneous CAPE forecasts, with a few-hour timing error changing CAPE values by one to two orders of magnitude. Based on these results and the corresponding discussion, forecasters should not blindly accept forecast values of CAPE. The models do appear to have skill in identifying the high CAPE events from the low CAPE events, thus alerting the forecaster to the potential of likely thunderstorm type (squall line vs multicell). However, the forecaster should be wary of using the model forecast CAPE to determine the thunderstorm potential to higher precision, given the models' drawbacks associated with surface temperature forecasts and the lack of cool downdrafts. The short-term forecasting method advised is to use the model forecast

soundings as a first guess for the current upper- and midlevel conditions, but to adjust the boundary layer temperature, dewpoint, and wind velocity to those values observed by surface observation stations and the NEXRAD velocity azimuth display (VAD) wind profile.

2) CONVECTIVE INHIBITION (CIN) FORECASTS

Figure 8 illustrates the forecast versus observed convective inhibition for each of the 18 convective events. Both models have little skill predicting CIN at 0000 UTC if the CIN exceeded 100 J kg^{-1} . The occurrences of underforecasted CIN can be partially attributed to the models' tendency to overheat the boundary layer, thus erroneously decreasing the inhibiting factor of warm

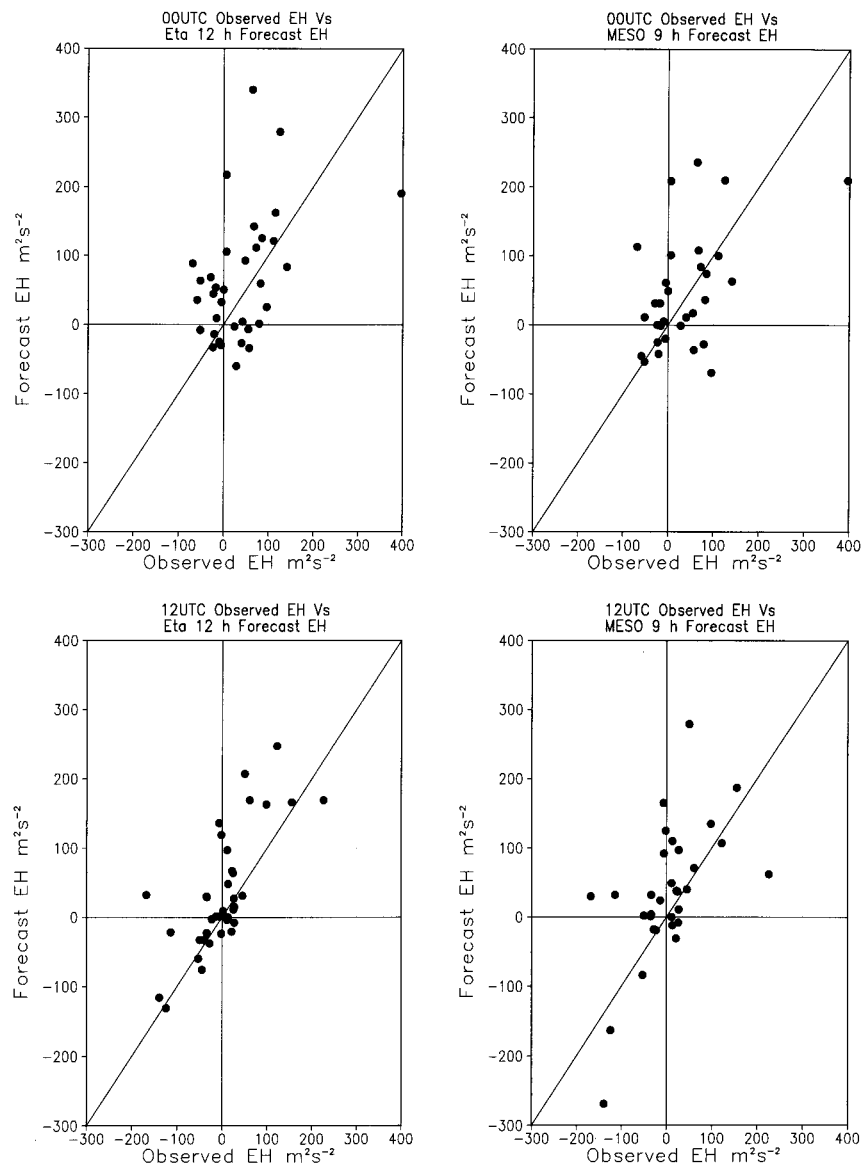


FIG. 9. Observed vs forecast environmental helicity (EH) for the Eta and MESO Models.

layers aloft. Again, the bias of underforecasting CIN was limited only by the tendency to underforecast dewpoints. In the few cases where the model overforecast convective inhibition at 0000 UTC, it was the result of errors in timing of frontal passages. If the model was too slow forecasting a cold frontal passage, the CIN would be overestimated; the opposite would be true for forecasts of erroneously early frontal passages.

In contrast to the forecasts valid at 0000 UTC, as the two lower panels in Fig. 8 illustrate, the models do appear to possess skill in forecasting CIN during the early morning (1200 UTC). Both the Eta and MESO appear biased toward underforecasting CIN, though the MESO has a smaller bias. One explanation for the greater skill at 1200 UTC is that the stable boundary layer

typically found during this time of day makes CIN forecasting sensitive to the conditions above the boundary layer. Thus, the CIN field at 1200 UTC is governed more by the synoptic-scale processes than by the mesoscale or storm-scale phenomena, as may be the case at 0000 UTC when convection is occurring or has already occurred in the atmosphere or the model. Second, 0000 UTC forecasts of surface equivalent potential temperature are warm biased (warm surface temperature bias exceeding cool dewpoint bias), leading to forecasts of CIN that are too low since the temperatures of ascending parcels are erroneously too warm. Based on these results, forecasters should be wary of using the model forecast soundings to anticipate the strength of convective inhibition. The intensity of the warm layer

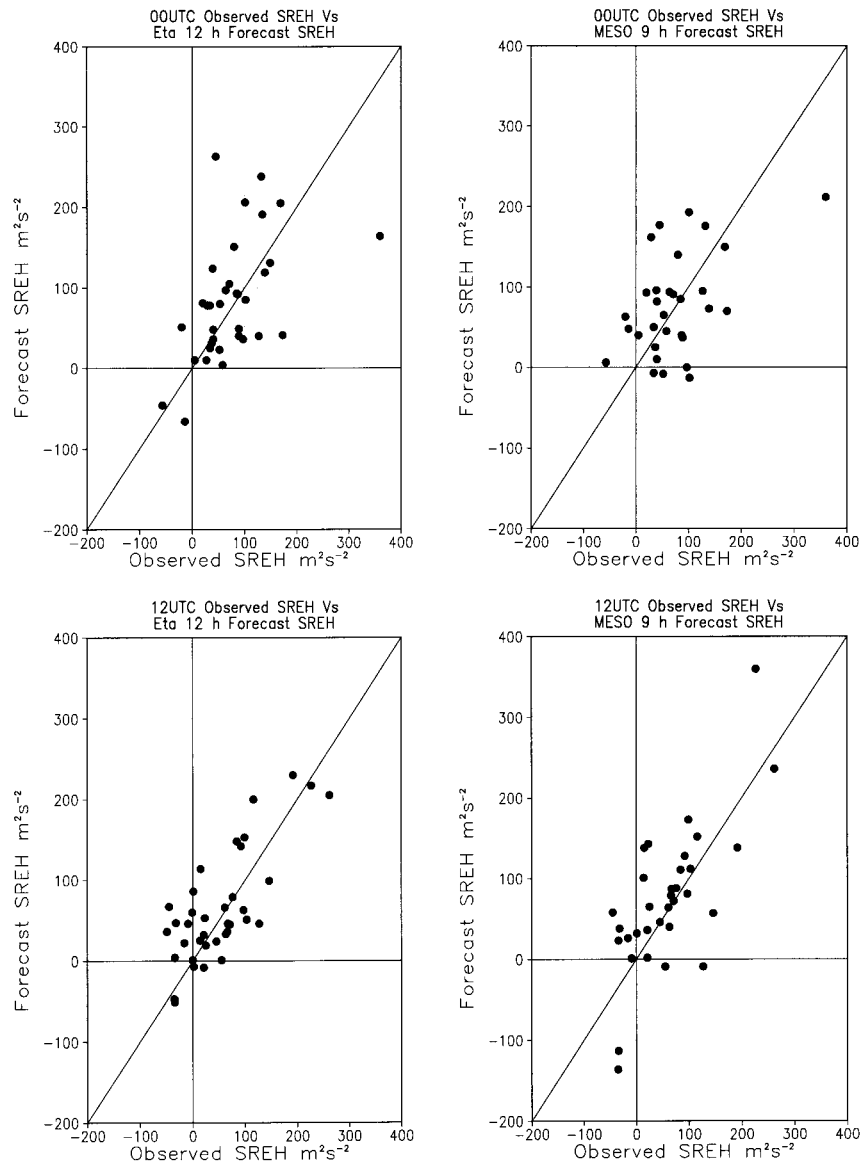


FIG. 10. Observed vs forecast storm-relative helicity (SREH) for the Eta and MESO Models.

aloft producing the inhibition was often correctly modeled, but the surface and boundary layer conditions at the same time were not as well forecasted.

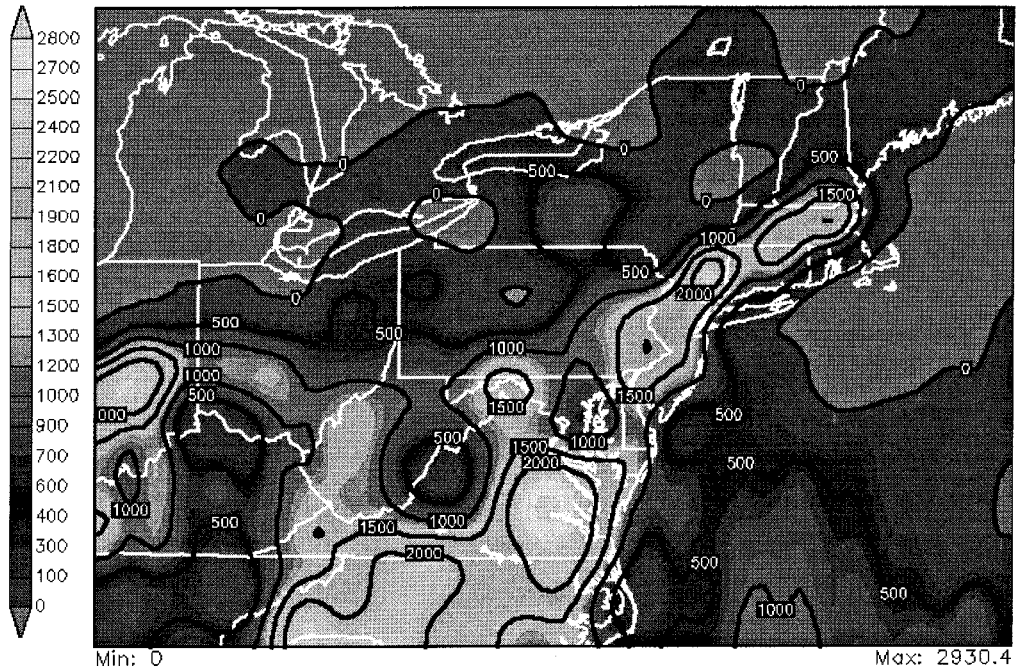
3) ENVIRONMENTAL AND STORM-RELATIVE HELICITY FORECASTS

During the 18 convective events, the environmental and storm-relative helicity were calculated for each observation–forecast pair for both the Eta and MESO Models (Figs. 9 and 10). Storm-relative helicity was calculated using the method described in Davies-Jones et al. (1990). Each model appears to have skill in predicting the environmental helicity at both 0000 and 1200 UTC, although there is a slight tendency to overpredict

it. Neither model appears to have significantly more skill than the other, which can probably be attributed to the high vertical resolution of both models. Surprisingly, Fig. 10 indicates that both models have more skill in predicting storm-relative helicity than environmental helicity, with slightly more skill at 1200 UTC than at 0000 UTC. However, in many cases during the 2-month evaluation period storm-relative helicity was increased on the mesoscale or storm scale due to storm–environment interaction that was not predicted by the model. This mesoscale or storm-scale increase of storm-relative helicity could be seen on the NEXRAD VAD wind profile. Examples of this are discussed in great detail in Pearce (1997). The models may provide for a reasonable prediction of the helicity field, but forecasters must be

(a)

Using 98052912 run eta forecast vertical temperature profile above surface
 Estimated Sfc-Based CAPE (J/kg) at 21Z/MAY 29,1998



(b)

Using 98052912 run eta forecast vertical temperature profile above surface
 2-Hr Lifted Index Change at 21Z/MAY 29,1998

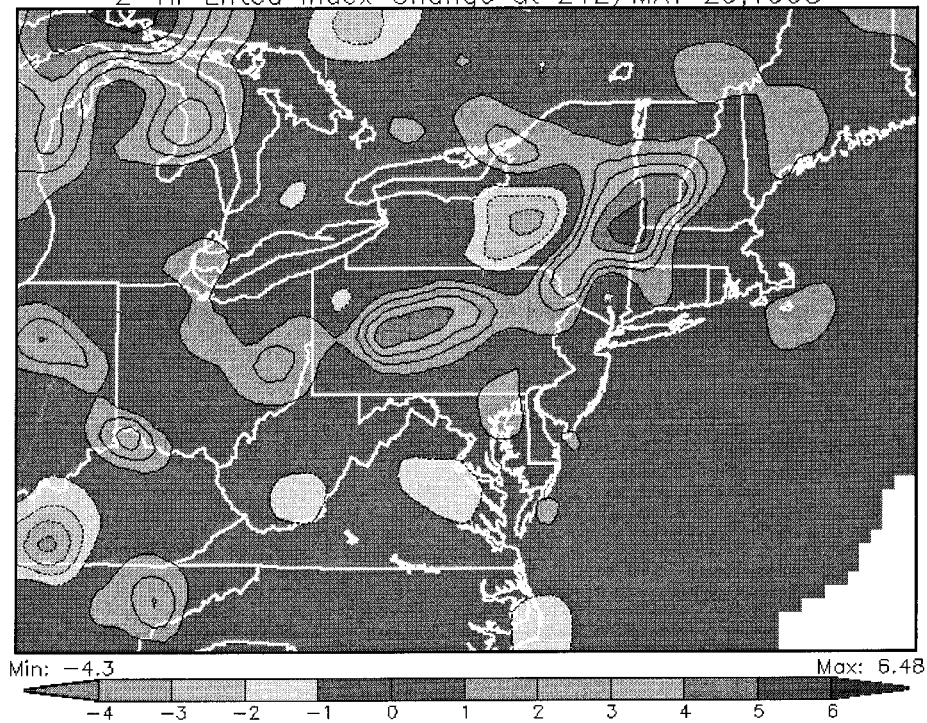


FIG. 11. (a) Example hourly model sounding-enhanced real-time analysis of CAPE. This particular analysis shows the moderately unstable air to the south of a southward moving cold front. (b) Example hourly model sounding enhanced real-time analysis of 2-h change in lifted index. Solid contour indicate positive values, while

alert to dramatically increased localized regions of higher storm-relative helicity resulting from storm–environment interaction (Pearce 1997).

During those days when the environmental or storm-relative helicity was underestimated, it was often due to the models' inability to veer boundary layer winds sufficiently in the presence of a weak southeasterly flow. Quite often during the evaluation period low-level winds (below 950 mb) would have a slight easterly component while the models would consistently forecast a slight westerly component. One possible explanation for this error is the overheating of the boundary layer previously discussed, which would tend to overmix and overdeepen the boundary layer depth, erroneously overmixing westerly momentum closer to the surface and reducing the ageostrophic fraction of the total wind at the surface. This 20°–40° forecast error in wind direction can dramatically influence the helicity values. Forecasters must be alert to such model forecast errors since such an error can create an environment less suitable to mesocyclone development than the real conditions may indicate. As noted previously, the parameterization errors leading to this bias have been fixed and, thus, we might expect that the errors in forecast low-level wind velocity should be reduced.

4) CONVECTIVE INITIATION FORECASTS

During weakly forced events, determining the locations and timing of thunderstorm development can be extremely difficult. If the synoptic or mesoscale dynamics are weak, determining the location of thunderstorm development hinges on determining the locations where surface heating is maximized, surface wind convergence is increased (through mesoscale, cloud, or topographic boundaries), and convective inhibition is first overcome. The timing of convective initiation (the point at which convective development is no longer inhibited by thermodynamic limitations) is typically given by the point at which the surface temperature and dewpoint are sufficiently high that rising parcels are able to just barely overcome a layer of CIN.

During the convective season of 1996, the forecasts of convective initiation by the Eta and MESO Models were examined, with a focus on weakly forced events. Observed convective initiation was determined when towering cumulus clouds were first indicated in surface reports. Model forecast convective initiation was said to occur when forecast CAPE exceeded forecast CIN. Overall, the results were fair, with convective development typically occurring within 1–3 h of the model forecast development (based on CAPE–CIN). There

were, however, numerous cases where timing errors were significant.

The greatest source of convective initiation timing error resulted from the tendency of both models to dissipate low-level stratus and fog too early, especially within and east of the Appalachian Mountains, as described previously. If the models removed the low-level stratus and fog too early, the boundary layer would be forecast to warm too quickly, and the convective temperature would be reached too early. This would erroneously start convection too early and too far east. In reality, the boundary layer would remain moist and slightly cooler east of the mountains if an easterly component of the low-level flow existed, while the boundary layer would warm much more quickly west of the mountains. Thus, convection would initiate too late in the east and too early the west. The tendency to initiate convection too early in weakly forced events was also supported by the models' tendencies to overforecast shelter temperatures. In cases where the environment was strongly forced, errors in convective timing were more often the result of erroneous forecast timing of synoptic features, such as fronts or jet streaks. Forecasters are advised to compare observed shelter temperatures to forecast shelter temperatures during the midmorning hours in weakly forced situations to determine if convection is likely to be initiated earlier or later than models may indicate.

5) REAL-TIME MODEL SOUNDING ENHANCED CONVECTIVE ANALYSES

In the United States, upper-air observations are usually taken several hours before and after convective events. For forecasters to estimate convective potential during the late morning or early afternoon, morning soundings are typically adjusted to a forecast surface temperature or an observed surface temperature. This process is fairly limiting, however, since upper-air observations are scarce and, on average, only one observation exists per forecast area. Therefore, mesoscale forecasting of convective potential is nearly impossible using modified observed upper-air soundings. The hourly model soundings alleviate, to a large degree, the limitations of using observed upper-air soundings. Given the hourly resolution, there exists a one-to-one relationship between a forecast model sounding and an hourly surface report. Further, numerous model soundings exist within a forecast area, giving the ability to diagnose the convective potential on the mesoscale.

During the convective season of 1997, hourly model soundings were used to derive mesoscale analyses of

←

dashed contours indicate negative values. The stabilization of the atmosphere behind the cold front is shown well in this analysis. Two-hour changes in lifted index of 4–6 exist immediately behind the front. In contrast, the region ahead of the front (Maine, Virginia, and Cape Cod) is continuing to destabilize as shelter temperatures rise.

hourly (real time) convective potential. Approximately 50 model soundings were objectively analyzed to produce a three-dimensional grid of forecast thermodynamic conditions for each forecast hour. The observed surface conditions (temperature and dewpoint) were then used to determine the lifting condensation level (LCL). Parcel traces from the LCL were then compared to the interpolated hourly model sounding atmospheric profile for that forecast hour. Using this method, estimates of CAPE (Fig. 11a), CIN, and lifted index (Fig. 11b) were calculated and made available to forecasters through a Web site (<http://hail.met.psu.edu/comet/automet2/automet2.html>). In addition, analyses indicating the 2-h changes in these fields were also made available, giving forecasters a sense of the trends in convective potential. Finally, standard surface analyses using only surface observations were also made available (temperature, dewpoint, equivalent potential temperature, pressure change, moisture convergence).

While the synthesis of hourly model soundings and surface observations alleviates the spatial and temporal resolution problems associated with observed rawinsonde soundings, it does introduce a larger source of error. Any model forecast errors will be translated to the model-enhanced convective fields. This potential for error is minimized, however, by constantly updating the upper-air objective analysis through use of only the latest available set of hourly model soundings. Therefore, typically the upper-air forecast data used in the convective fields are no more than 6–9 h old. Forecasters at the National Weather Service Office in State College have expressed a high degree of satisfaction with the convective analyses. They have found the fields to be useful in anticipating the formation and dissipation of thunderstorms, mesoscale convective complexes, convective trends, mesoscale boundaries, and areas of moisture convergence.

4. Concluding summary and discussion

This research examined the utility of high-resolution numerical model vertical profiles to forecast mesoscale phenomena. The vertical profiles were visualized in four formats: time–height cross sections, time series histograms, skew T – $\log p$ animations, and distance–height cross sections. The ability of the NGM, Eta, and MESO Models to forecast warm-season mesoscale phenomena was examined. The most significant forecast errors were in the forecast shelter-level temperatures, which produced a cascade of impacts on the forecast accuracy of other mesoscale phenomena, as described below.

The 2-m temperature forecast by the Eta and MESO Models can be used as a forecast for the local surface high and low temperatures. Overall, there exists a warm bias in the 2-m temperature during the daytime and a cold bias at night. This bias is maximized during clear skies and light winds. The warm daytime bias also overdeepens the boundary layer, lowering surface dewpoints

too greatly and overmixing westerly momentum to the surface. Improvements in the model radiation parameterizations in early 1997 have resulted in a decrease in the magnitude of the 2-m temperature bias. Preliminary indications are that a bias still remains, with a slight cold bias (1° – 4° C) during the daytime and a slight warm bias at night (1° – 2° C).

Using the boundary layer relative humidity, the NGM, Eta, and MESO all showed some degree of skill in forecasting fog. The Eta and MESO produced fog at lower empirical relative humidity thresholds than the NGM: 92%, 93%, and 95%, respectively. There was limited to little skill in predicting localized ground fogs, however. Using the experimental fog product, there did exist skill in forecasting the timing of fog burnoff. However, the Eta and MESO had a tendency to burn off fog too quickly in the morning, a consequence of the PBL warm temperature bias. This tendency to decrease low-level relative humidity too quickly was also seen in the models' tendency to remove morning stratus clouds too quickly. However, by examining the vertical and temporal distribution of relative humidity in the boundary layer, the forecaster is able to overcome the models' fog–stratus removal and associated warm PBL biases. Once this bias was accounted for, skill did exist to forecast the removal of early morning stratus.

Modest skill existed in forecasting convection using the Eta and MESO Models. The greatest detriments to the models' ability to forecast convection were twofold: first, the PBL warm bias produced overestimates of CAPE and underestimates of CIN. In addition, the increased boundary layer depth produced an overmixing of westerly momentum to the surface, which greatly impacted helicity forecasts. Second, the Betts–Miller convective parameterization scheme is not well suited for forecasting convection in the middle latitudes. The scheme is unable to produce cool downdrafts, cannot produce realistic rainfall rates, and is unable to simulate elevated convection.

The most promising results of this study appeared to be in forecasting turbulence. The increased accuracy of the results may partially be attributed to the relative independence of turbulence outside the boundary layer from the PBL thermal bias. Forecast Richardson number values of 0.25 or less can be associated with turbulence. Further, values of Richardson number of less than 1 were highly correlated with a high probability of turbulence. These results show the potential for artificial intelligence applications to generate computer-worded aircraft routing forecasts.

Significant spatial errors exist between the Eta–MESO Model sounding location and the observation station location. The consequences of this displacement could be seen in forecasts of all warm-season phenomena: surface temperature, precipitation, and fog. The consequences of this displacement are greatest near coastlines and in regions of large terrain gradients. Future studies involving verification of convective precip-

itation events will likely require interpolation between model sounding points.

Given the high-resolution and locale-specific forecasts provided through these model profiles, research could be applied to produce artificial intelligence applications that would generate worded forecasts to accompany the visualization tools shown here. Development of experimental products during this research was an integral portion of the success of the project. Using these products, forecasters now have nonofficial guidance for the occurrence of fog and the threat of turbulence to pilots, 1–2 days in advance. These products have shown variable skill and as forecasters become more familiar with the output produced by the products, their utility will only increase with time. However, it should be emphasized that the criteria found during this research are not necessarily directly applicable to other regions of the country. Further refinement of empirical thresholds for other (non-Pennsylvania) locales will likely be necessary. In addition, the criteria developed during this research may not perform as well in the future as changes to model physics, parameterizations, and resolution occur.

Finally, during the course of this research project, NCEP was provided with the preliminary findings and they added model sounding sites to expand and improve the overall study. This led to NCEP's continued refinements and improvements in the parameterization schemes in the Eta and MESO during the course of this project. This interactive relationship facilitated the improvement of the model profiles' usefulness and accuracy as the project continued.

Acknowledgments. The authors gratefully acknowledge the support of the National Weather Service and COMET for supporting this research through a NWS–COMET Fellowship. In particular, funds were provided by the University Corporation for Atmospheric Research (UCAR) Subawards UCAR S95-59695 and UCAR S96-75664 and pursuant to the National Oceanic and Atmospheric Administration Awards NA37WD0018-01V and NA57GP0576. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA, its subagencies, or UCAR.

Numerous scientists, students, and forecasters provided invaluable guidance and support during the two years of research, without whom this research could not have been as complete. Keith Brill and Dr. Geoff DiMego of NCEP provided the dataset on which this research was supported. Their efforts to constantly improve the quality, completeness, and timeliness of the dataset are greatly appreciated. The efforts to modify the scope of the dataset to meet the needs of this research

are also appreciated, including the addition of several model sounding stations in the Pennsylvania region. Finally, the authors thank the Center for Ocean–Land–Atmosphere Studies (COLA) of the University of Maryland for the use of the GrADS software package. Additionally, we would like to thank Dr. Mike Fiorino of the Lawrence Livermore National Laboratory for his help with the package during the early stages of this research, including the development of the skew T –log p routine.

REFERENCES

- Black, T. L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting*, **9**, 265–278.
- , D. G. Deaven, and G. DiMego, 1993: The step-mountain eta coordinate model: 80 km early version and objective verifications. NWS Tech. Procedures Bull. 412, 31 pp. [Available from National Weather Service, Office of Meteorology, 1325 East–West Highway, Silver Spring, MD 20910.]
- Dallavalle, J. P., J. S. Jensenius Jr., and S. A. Gilbert, 1992: NGM-based MOS guidance—The FOUS14/FWC message. NWS Tech. Procedures Bull. 408, 19 pp. [Available from National Weather Service, Office of Meteorology, 1325 East–West Highway, Silver Spring, MD 20910.]
- Davies-Jones, R., D. W. Burgess, and M. Foster, 1990: Test of helicity as a tornado forecast parameter. Preprints, *16th Conf. Severe Local Storms*, Kananaskis Park, AB, Canada, Amer. Meteor. Soc., 588–592.
- desJardins, M. L., and R. A. Petersen, 1985: GEMPAK: A meteorological system for research and education. Preprints, *First Conf. on Interactive Information and Processing Systems*, Los Angeles, CA, Amer. Meteor. Soc., 313–319.
- Gaza, R. S., and L. F. Bosart, 1985: The Kansas City severe weather event of 4 June 1979. *Mon. Wea. Rev.*, **113**, 1320–1330.
- Hart, R. E., 1997: Forecasting studies using hourly model-generated soundings. M.S. thesis, Department of Meteorology, The Pennsylvania State University, University Park, PA, 166 pp. [Available from Dept. of Meteorology, The Pennsylvania State University, 503 Walker Bldg., University Park, PA 16802.]
- Hoke, J. E., N. A. Phillips, G. J. DiMego, J. J. Tucillo, and J. Sela, 1989: The Regional Analysis and Forecast System of the National Meteorological Center. *Wea. Forecasting*, **4**, 323–334.
- Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945.
- Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, **7**, 588–612.
- Niziol, T. A., and E. A. Mahoney, 1997: The use of high resolution hourly soundings for the prediction of lake effect snow. Preprints, *13th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 92–95.
- Pearce, M. L., 1997: Non-classic and weakly forced convective events: A forecasting challenge for the dominant form of severe weather in the mid-Atlantic region of the United States. M.S. thesis, Department of Meteorology, The Pennsylvania State University, University Park, PA, 122 pp. [Available from Dept. of Meteorology, The Pennsylvania State University, 503 Walker Bldg., University Park, PA 16802.]