Short-Range (0–48 h) Numerical Prediction of Convective Occurrence, Mode, and Location

MICHAEL A. FOWLE* AND PAUL J. ROEBBER

Atmospheric Sciences Group, Department of Mathematical Sciences, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin

(Manuscript received 16 July 2002, in final form 14 January 2003)

ABSTRACT

A verification of high-resolution (6-km grid spacing) short-range (0–48 h) numerical model forecasts of warm-season convective occurrence, mode, and location was conducted over the Lake Michigan region. All available days from 5 April through 20 September 1999 were evaluated using 0.5° base reflectivity and accumulated precipitation products from the national radar network and the day-1 (0–24 h) and day-2 (24–48 h) forecasts from a quasi-operational version of the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5). Contingency measures show forecast skill for convective occurrence is high, with day-1 (day 2) equitable threat scores and Kuipers skill scores (KSS) of 0.69 (0.60) and 0.84 (0.75), respectively. Forecast skill in predicting convective mode (defined as linear, multicellular, or isolated) is also high, with KSS of 0.91 (0.86) for day 1 (day 2). Median timing errors for convective initiation/dissipation were within 2.5 h for all modes of convection at both forecast ranges. Forecasts of the areal coverage of the 24-h accumulated precipitation in convective events exhibited skill comparable to the lower-resolution, operational models, with median threat scores at day 1 (day 2) of 0.21 (0.24). When small displacements (less than 85 km) in the precipitation pattern were taken into account, threat scores increased to as high as 0.44 for the most organized convective modes. The implications of these results for the use of mesoscale models in operational forecasting are discussed.

1. Introduction

Operational numerical weather prediction (NWP) models have shown relatively steady increases in skill in synoptic-scale forecasts of the atmospheric circulation over the past several decades (e.g., Fig. 1 of Roebber and Bosart 1998). One prominent example of these improvements is in the forecasts of the development and evolution of extratropical cyclones (e.g., Sanders 1987; Oravec and Grumm 1993). Despite these advances, the quantitative precipitation forecast (QPF) problem continues to pose a significant operational forecast challenge (e.g., Olson et al. 1995; Roebber and Bosart 1998; Fritsch et al. 1998), most particularly in the warm season {e.g., Olson et al. (1995) show average warm-season threat scores for 25-mm forecasts for the period of 1984–93 of less than 0.15, where perfect forecasts equal 1.00 [more recent statistics, posted on the National Centers for Environmental Prediction (NCEP) Web site at http://www.hpc.ncep.noaa.gov/html/hpcverif.shtml, reveal similar levels of skill]}. The problem of warm-season QPF is tied to the broader challenges associated with convective forecasting.

Numerical guidance on the problems of convective initiation (occurrence, timing, and location) and the resulting precipitation amounts remains poor, and guidance on convective mode and associated sensible weather is generally not available. Because the dynamic and thermodynamic aspects of convective episodes occur on the mesoscale, it is perhaps unsurprising that guidance on these problems is unsatisfactory. Errors can arise in at least three ways. First, it is recognized that forecasts of precipitation can be sensitive to initial condition uncertainty. Du et al. (1997) found, for a case of rapid cyclogenesis, that initial condition uncertainty led to variations in 6-h convective precipitation totals of as much as 75 mm in synoptic-scale model simulations. Roebber and Reuter (2002) showed that the amounts of convective precipitation that arise from synoptically and thermodynamically similar events can vary considerably, as a result of differences in the flow details that trigger the convection (in their case, coupling between upper- and lower-level jets). Second, because the operational NWP models lack the capability to resolve convection explicitly, warm-season QPFs are derived partly from subgrid-scale precipitation processes that are represented using cumulus parameterization

* Current affiliation: Iowa Department of Natural Resources, Urbandale, Iowa.

Corresponding author address: Paul J. Roebber, Dept. of Mathematical Sciences, University of Wisconsin—Milwaukee, 3200 N. Cramer Ave., Milwaukee, WI 53211.
E-mail: roebber@uwm.edu

schemes. The challenges of this approach are well known (e.g., Emanuel and Raymond 1993; Wang and Seaman 1997; Liu et al. 2001). Further, the method is (by definition) incapable of providing guidance on convective details (mechanisms of convective initiation, determinations of convective mode, and resultant sensible weather). Third, our understanding of precipitation processes in general, and cloud microphysical issues in particular, remains limited, owing in part to a paucity of data at those scales.

These three sources of error suggest the following remedies: countering limitations imposed by the lack of data and concomitant understanding at the cloud, meso-, and synoptic scales requires increased observations; a possible means of countering limitations imposed by subgrid-scale precipitation processes is through the explicit prediction of convection (subject to the limits imposed by model physics errors). Anthes (1986) hypothesized that, in at least some regions and at some times, the mesoscale may be dominated by synoptic-scale processes and would therefore be relatively predictable [see Gall and Shapiro (2000) for further discussion of this point]. Hence, it is clear that explicit prediction cannot be completely separated from the need for increased observations, because the efficacy of the approach will depend on accurate simulation of the range of scales governing the convection. Stensrud et al. (2000) and Stensrud and Weiss (2002) have shown that model physics errors are also an important contributor to forecast uncertainty for convective events. Conclusions drawn concerning improvements that result from explicit representation of convection using current observational networks and model physics must be considered as a conservative estimate of what may be possible.

Nonetheless, there is much to be learned from the immediate study of this problem. With advances in computer power and numerical modeling of physical processes, it is now *possible* to run much higher resolution models in operational settings. The question that remains to be answered is whether this is an effective use of computer resources relative to other approaches, such as coarse-resolution ensemble forecasts that are combined with techniques to infer details such as convective mode (e.g., Stensrud et al. 1997). To what extent can the forecast problems of convective occurrence, mode, and location be addressed at these forecast ranges, given data limitations? Recently, hindcast (Liu et al. 1997; Bernardet and Cotton 1998; Nachamkin and Cotton 2000; Finley et al. 2001; Romero et al. 2001; Bélair and Mailhot 2001) and forecast model (Roebber et al. 2002) simulations of specific events have suggested that considerable insight into convective details, including convective mode, can be obtained from explicit model representations.

The specific objective of this work is to document the capability of a high-resolution (6-km grid spacing), quasi-operational NWP model to predict convective occurrence, mode, and location in the 0–48-h forecast range for a large sample and variety of convective episodes. Previous studies have raised the possibility that high-resolution forecasts may be most useful in developing a conceptual understanding of a situation rather than providing explicit predictions of individual storms (Brooks et al. 1992; Roebber et al. 2002). This issue will be discussed in the context of the findings reported in this paper.

The organization is as follows. Section 2 presents the model description, data, and analysis procedures. In section 3, verification results for convective occurrence, mode, and location are provided. A summary, including some discussion of the use of models as a conceptual basis for forecasting, is presented in section 4.

2. Model description, data, and analysis procedures

a. Model description

The University of Wisconsin—Milwaukee (UWM) real-time forecast system is based on the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5), a nonhydrostatic, primitive equation model (Dudhia 1993; Grell et al. 1994). The model is launched once daily from a cold start, utilizing the 0000 UTC forecast cycle of the NCEP Aviation Model (AVN) run (Petersen and Stackpole 1989) to provide initial and lateral boundary conditions. A key point is that no data in addition to the AVN output (at T126 or ~105-km horizontal resolution) were used to initialize the MM5 on any of the domains. The MM5 forecasts are conducted in triply nested, two-way interactive mode (Fig. 1), such that conditions in an inner domain feed back to the coarse domain, and vice versa, with matching at the nest boundary (Zhang et al. 1986). The outermost domain (D1), with 54-km horizontal grid spacing, is designed to capture the synoptic-scale features that affect the day-to-day sensible weather. The model is nested through domain two (D2, 18 km) to domain three (D3, 6-km), with these latter domains designed to resolve mesoscale features. The verifications in this study will be conducted on D3, because features associated with warm-season convection are likely to be resolved best at this grid scale.

An explicit moisture scheme with prognostic equations for cloud water, ice, rainwater, and snow (Reisner et al. 1998) was employed in all domains for grid-resolvable precipitation. The Kain–Fritsch cumulus parameterization scheme (Kain and Fritsch 1993) was used in the two outermost domains (D1, D2) while in D3 no convective parameterization is used, under the assumption that at 6-km grid spacing convective features can be broadly resolved. Radiative processes were handled using a cloud-radiation scheme in which diurnally varying shortwave and longwave radiative fluxes interact with explicit cloud and clear air, and the surface fluxes

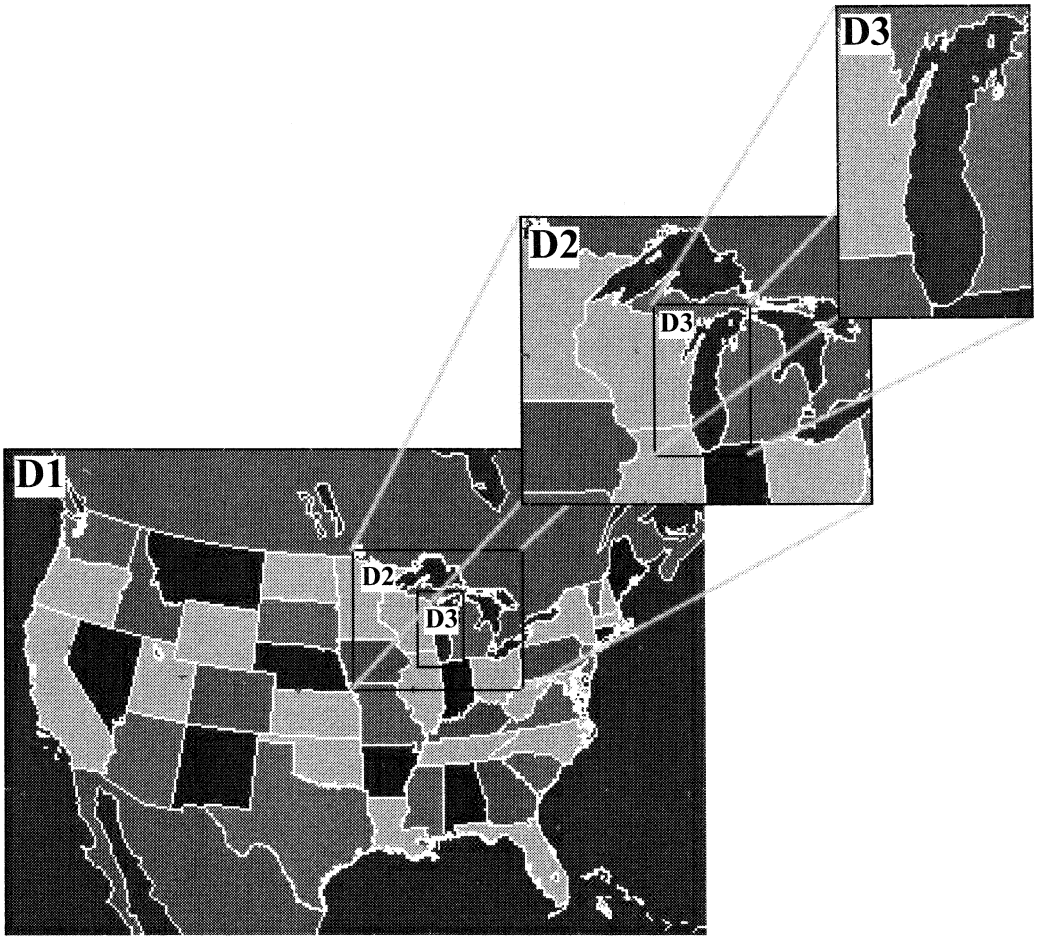

FIG. 1. MM5 nested model domains. Horizontal grid spacing is 54, 18, and 6 km in D1, D2, and D3, respectively.

were used in the ground energy budget calculations (Dudhia 1989). The planetary boundary layer was modeled using the high-resolution Blackadar scheme (Blackadar 1979; Zhang and Anthes 1982) coupled with a five-layer soil model (Dudhia 1996). Standard procedures for initializing the soil temperature profile and soil moisture were used (linear temperature variation with depth, based on ground and substrate temperatures at the initial time, and climatological lookup tables for moisture, respectively). The model had 23 vertical sigma levels, with a relative concentration at the lowest levels to resolve planetary boundary layer structures important to convective initiation.

b. Observational data

The forecast verification was conducted from 5 April to 20 September 1999 (Table 1). The observational data used to verify the occurrence of convection were primarily from the high-resolution Weather Surveillance Radar-1988 Doppler (WSR-88D). The WSR-88D data were acquired from the National Climatic Data Center's (NCDC) National Mosaic Reflectivity Images online archive Web site (http://www4.ncdc.noaa.gov/cgi-win/wwcgi.dll?WWNEXRAD~Images2), which provided hourly reflectivity data for the period of interest. On days when the convection was less pronounced (i.e., small areal and/or temporal coverage), hard copies of 0.5° base reflectivity data from the Green Bay WSR-88D site were examined. The temporal frequency of these images ranged from every 30 min to every 2 h, depending on the needs of the case. The primary advantage of the hard copies was the increased ability to discern individual cells in comparison with the coarse resolution of the NCDC mosaic.

Although the WSR-88D data are used as the observational basis, it should not be assumed that these data are without problems. Many factors, such as convective position relative to the radar location and radar downtime, can lead to errors in these analyses. To aid further in storm documentation, routine surface observations of precipitation were included. It is important to note that surface weather observations alone may underestimate the occurrence of precipitation given the nonuniform

TABLE 1. Observed convective days examined in the study in day-1 (0-24 h) and day-2 (24–48 h) forecast ranges for the period of 5 Apr–20 Sep 1999. An X indicates that the day was examined in the indicated forecast range.

Date	Day 1	Day 2
6 Apr	X	X
7 Apr	X	X
10 Apr	X	
11 Apr	X	X
15 Apr	X	X
16 Apr		X
19 Apr		X
20 Apr		X
21 Apr	X	
22 Apr	X	X
23 Apr	X	X
28 Apr	X	
5 May	X	X
6 May	X	X
7 May		X
13 May		X
15 May		X
16 May	X	
20 May	X	X
21 May	X	X
22 May	X	X
23 May		X
24 May		X
31 May	X	
1 Jun		X
2 Jun	X	X
4 Jun	X	X
5 Jun	X	X
6 Jun	X	X
7 Jun	X	
8 Jun	X	
9 Jun	X	
10 Jun	X	
12 Jun	X	
13 Jun	X	X
14 Jun	X	X
16 Jun	X	X
17 Jun	X	X
20 Jun		X
22 Jun	X	
23 Jun	X	X
24 Jun	X	X
27 Jun	X	X
28 Jun	X	X
29 Jun	X	X
1 Jul	X	X
2 Jul	X	X
3 Jul	X	X
6 Jul	X	X
8 Jul	X	X
9 Jul	X	X
13 Jul		X
14 Jul	X	
15 Jul	X	
16 Jul	X	
17 Jul	X	X
19 Jul	X	X
3 Aug	X	
4 Aug	X	
6 Aug		X
7 Aug	X	X
9 Aug	X	X
10 Aug	X	X
11 Aug		X
12 Aug	X	X
13 Aug	X	
16 Aug	X	X
18 Aug	X	X
19 Aug	X	X
21 Aug	X	
22 Aug	X	
23 Aug	X	X
24 Aug	X	X
25 Aug	X	X
5 Sep	X	X
6 Sep		X
8 Sep	X	X
12 Sep	X	X
13 Sep	X	X
19 Sep	X	X
20 Sep	X	X

spatial distribution of some modes of precipitation and limited station density. As a result, these data were used only as a secondary confirmation tool.

The observed areal coverage of precipitation was derived from the NCEP/Office of Hydrology (OH) hourly, multisensor National Precipitation Analysis (NPA). This analysis merges 3000 automated, hourly rain gauge observations in the contiguous 48 states with hourly digital-precipitation-data radar estimates. The data analysis routines, including the bias correction of the radar estimates using the gauge data [developed by OH and executed regionally at National Weather Service (NWS) River Forecast Centers], have been adapted by NCEP to a national 4-km grid. In this study, the multisensor analysis (gauge and unbiased radar) was utilized.

c. Verification procedure

For the period of 5 April–20 September 1999, the verification of convective precipitation was conducted for all days on which joint model and observed data were available. Accounting for missing data, this total comprised 143 (141) days at the 0–24-h (24–48 h) forecast range. A multistep approach to the verification was conducted and is described below. All verifications were conducted on D3 (6-km grid spacing).

1) CONVECTIVE OCCURRENCE

The first stage of the verification focused on the daily occurrence of precipitation. This entailed a binary determination of the existence of convection in the observations and model forecasts for each day of the period. This structure was chosen so that model performance could be evaluated through the use of contingency tables (see section 3). Convective echoes were observationally defined as those with reflectivity factors greater than 40 dBZ, a definition corresponding to cri-

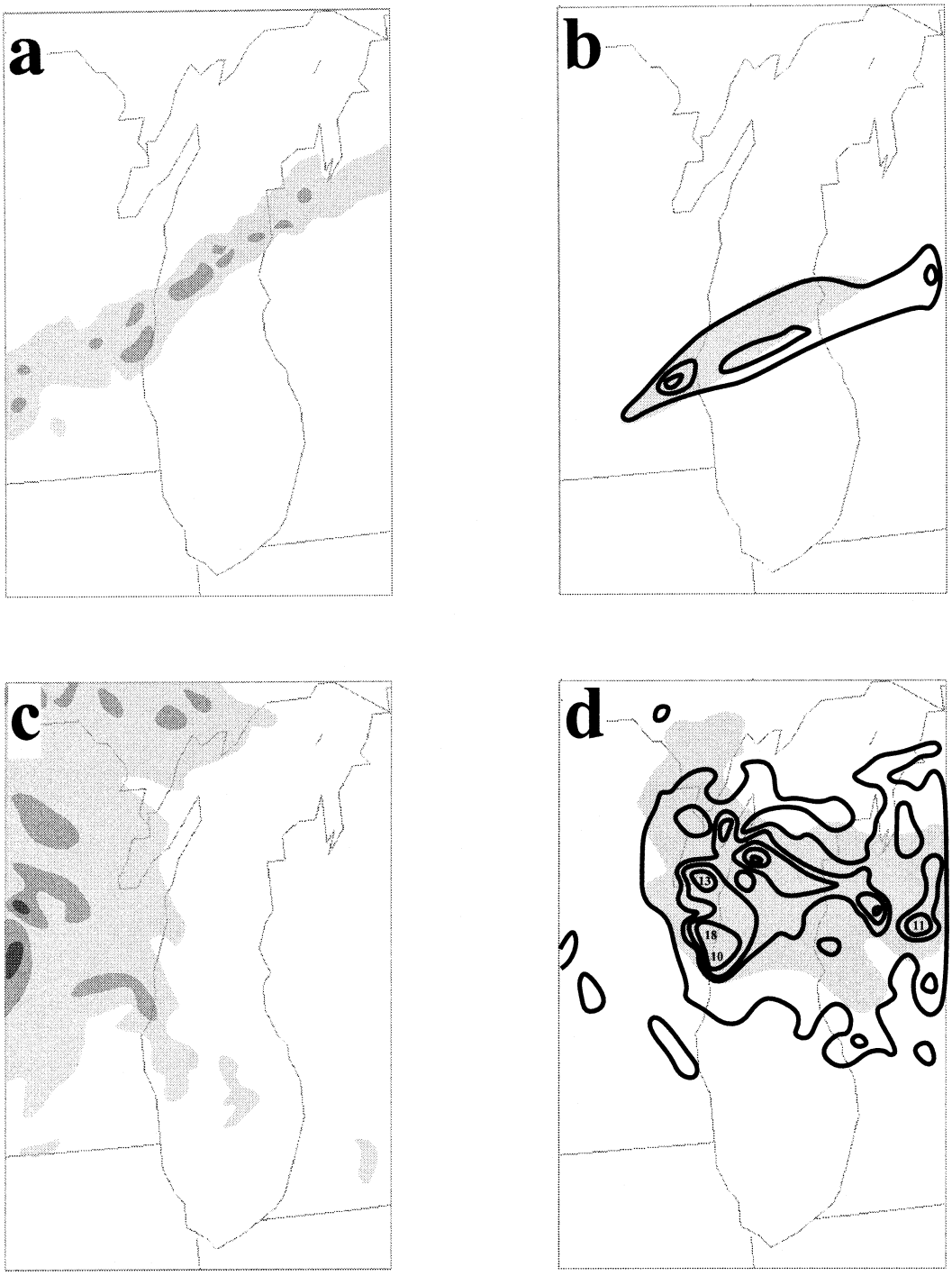

FIG. 2. Representative (a), (c), (e) observed and (b), (d), (f) corresponding model forecast convection, identified following the criteria discussed in section 2c. For the observed cases, radar reflectivity is shown (25–30 dBZ, light gray shading; 40–45 dBZ, dark gray shading; >50 dBZ, black shading). For the corresponding model cases, upward vertical velocity (bold contours, 0.6 m s^{-1} interval, from 0.3 to 2.1 m s^{-1}; some maximum values given in meters per second) and rainwater (shaded, greater than 0.1 g kg^{-1} at the 0.87 sigma level) are shown. The cases are (a), (b) a linear event at 1400 UTC 16 Aug 1999, (c), (d) a multicellular event at 1400 UTC 3 Jul 1999, and (e), (f) an isolated event at 2200 UTC 21 Aug 1999.

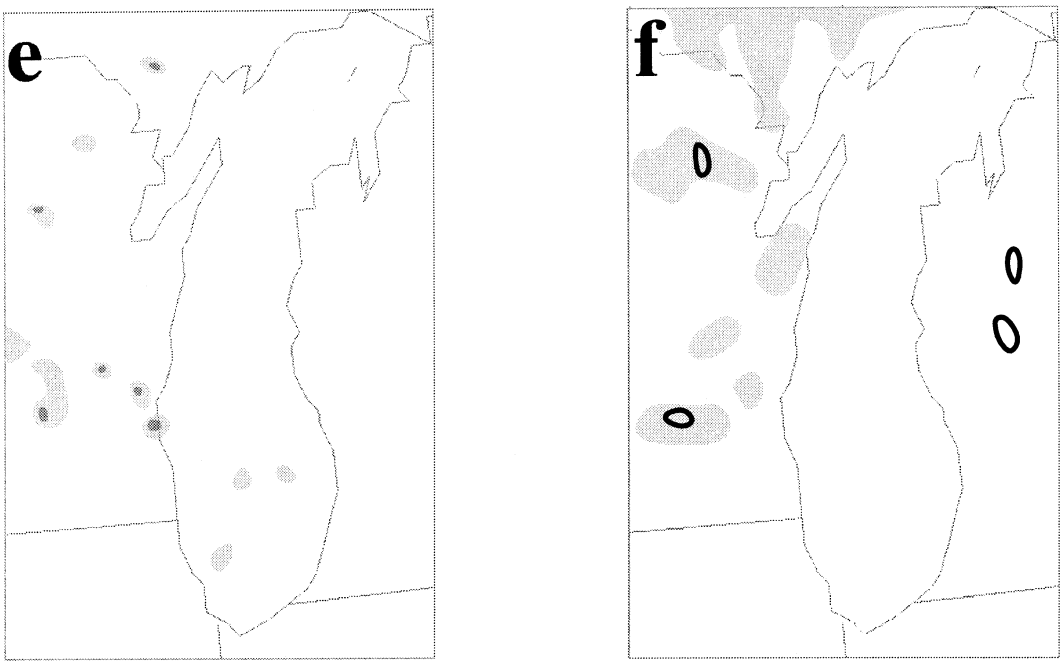

FIG. 2. (*Continued*)

teria used in previous studies (e.g., Steiner et al. 1995; Geerts 1998).

Model convection was judged on the criteria of maximum upward vertical velocity (UVV) in conjunction with the production of rainwater. The threshold value of UVV utilized to identify convection was taken to be greater than 0.3 m s^{-1}. This value of UVV is low in comparison with observed values of approximately 2–5 m s^{-1} [with peak values in excess of 40 m s^{-1} in some severe storms; e.g., Bluestein et al. (1988)]. However, this definition is constructed based on the understanding that at a grid spacing of 6 km the magnitude of convective motions is not fully resolved.

An examination of model UVV for 30 convective days in the two interior domains (D2 at 18-km and D3 at 6-km grid spacing) was conducted to evaluate the validity of the 0.3 m s^{-1} UVV threshold. In each of the 30 cases, one hour was isolated, and the maximum UVV at a specific latitude and longitude was documented for the two model domains. Median (50th percentile) results show the magnitude of the maximum UVV to be a factor of 3.0 greater in the 6-km domain (D3) than in the 18-km domain (D2). If this ''rule of three'' persists to higher resolutions, at which convective motions are more fully simulated (e.g., a decrease in grid spacing from 6 to 2 km), then the minimum threshold value of UVV used in the study is consistent with weak observed convective motions [e.g., Houze (1993), who defines convection as having vertical velocities greater than the 1 m s^{-1} fall speed of snow). An a posteriori justification of this threshold is provided by the fact that overforecasts of convection did not occur and that the forecasts of convection are highly correlated with observed occurrences as confirmed by radar. Examples of three observed and modeled convective cases, identified using these criteria, are presented in Fig. 2. The characteristics of the observed convection are well represented in the model forecasts. Further details of this correspondence are presented in section 2c(2).

2) CONVECTIVE MODE

The second stage was the verification of convective mode. Observed storms were categorized as linear, multicellular, or isolated (pulse storms). These categorizations were determined according to the most organized stage of the storm lifespan (with linear considered to be the most organized and isolated the least), based upon the reflectivity signature (e.g., Weisman and Klemp 1986; Hane 1986).

A strict definition of what constitutes linear convection does not exist. Bluestein (1993, p. 520) states: ''The radar echo associated with a precipitation 'band' is at least 5 times long as wide, at least 5–10 km wide, and persists for hours.'' Utilizing this guidance, a minimum areal criterion of 500 km^2 was defined (10-km width, 50-km length). Parker and Johnson (2000) state: ''Conglomerates of convective storms are often organized on the mesoscale, behaving as long lived (greater than 3 h) discrete entities called mesoscale convective systems (MCSs).'' This suggests a minimum persistence criterion of 3 h. Because Bluestein (1993) does not formally distinguish between stratiform and convective echoes (e.g., Houze 1993; Geerts 1998), in this study, in which

only convective echos are considered (>40 dBZ), a length:width ratio of 3:1 was applied. Taken together, these criteria result in the identification of linear convection if the radar reflectivity > 40 dBZ exhibits a length:width ratio of at least 3:1, persists for at least 3 h, and has an areal coverage of at least 500 km^2.

Multicellular convection was defined by a reflectivity area greater than 40 dBZ with a length:width ratio less than 3:1 that persists for at least 3 h and has an areal coverage of at least 500 km^2. Studies have shown (e.g., Houze et al. 1989; Loehrer and Johnson 1995) that mesoscale convective systems often exhibit a linearity associated with a localized region of very high reflectivity surrounded by much larger regions of slightly lower reflectivity. In these cases, the surrounding areas were included in the total reflectivity shape of the system.

The isolated storm mode was defined by a reflectivity area greater than 40 dBZ that had a spatial coverage of less than 500 km^2. In the study period, no cases meeting this definition persisted for more than 3 h. Because of the small areal coverage and brief life span, the isolated mode was the most difficult to identify. In these less distinct cases, the rain gauge analysis and the hard copy WSR-88D reflectivity images were used in the identification process.

Convective mode in MM5 was characterized using a method similar to that of the observations, except that, for identified model convection (UVV greater than 0.3 m s^{-1}), rainwater was used as a proxy for reflectivity. For a few events in the study (less than 5%), subjective judgments were made with respect to the relative spatial area of the rainwater shield and relative coalescence between convective cells to discriminate between multicellular and isolated convection. For instance, possible mergers between individual updrafts were monitored closely to indicate a further evolution of convective cells. Examples of observed and modeled convective modes, identified using these criteria, are shown in Fig. 2. The qualitative correspondence between the observed linear (Fig. 2a), multicellular (Fig. 2c), and isolated (Fig. 2e) convection and the respective forecast modes (Figs. 2b,d, and f, respectively) is clear.

3) CONVECTIVE LOCATION

The final verification stage examined the areal coverage of convective precipitation. This was accomplished for the observations by interpolating the 4-km NCEP blended precipitation to the model grid in binary form (yes–no); the model area was directly obtained from the model forecast precipitation. A complete description of this task will be presented in section 3c.

3. Convective verification results

a. Convective occurrence

The verification of convective occurrence is based upon a contingency table approach (Wilks 1995) in which each element of the table holds the number of occurrences in which the atmosphere and the model did or did not exhibit convection over a 24-h period (Fig. 3). Observed convection is defined as a 24-h period (0000–2359 UTC) in which radar reflectivities document at least one convective cell with a maximum reflectivity greater than 40 dBZ for a minimum of 1 h, anywhere within the 6-km domain (D3). The definition of modeled convection was met when UVV exceeded 0.3 m s^{-1} and was accompanied by the production of rainwater, for a minimum of 1 h, anywhere within the region of the 6-km domain (D3). The four elements of the contingency table are then defined as follows: a = convection both observed and forecast, b = convection forecast but not observed, c = convection observed but not forecast, and d = convection neither observed nor forecast.

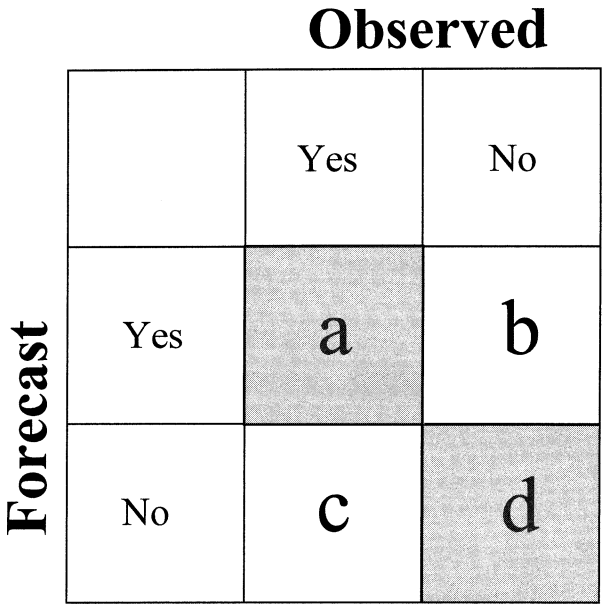

FIG. 3. A 2 × 2 contingency table structure. Shaded boxes denote correct forecasts. See text for details.

The following statistical measures were utilized: probability of detection (POD; the fraction of observed events correctly forecast), false-alarm ratio (FAR; the fraction of event forecasts that do not verify), bias (the ratio of forecast to observed events), canonical threat score (TS; fraction of events correctly forecast minus the nonevent cases), equitable threat score (ETS; like the TS but accounting for the number of forecasts expected to be correct by chance), and the Kuipers skill score (KSS, also known as the true skill statistic; a skill measure relative to a baseline of random forecasts, climatological values, or constant forecasts of any kind). These measures are defined as follows:

$$\mathrm{POD} = \frac{a}{a + c}, \tag{1}$$

TABLE 2. MM5 forecast performance for day 1 (0–24 h) and day 2 (24–48 h) for convective occurrence within D3.

		Observed	
		Yes	No
Raw totals			
MM5 0–24-h forecast	Yes	67	0
	No	13	63
MM5 24–48-h forecast	Yes	62	4
	No	14	61

Summary measures of forecast performance

Forecast measure	0–24-h forecast	24-48-h forecast
POD	0.84	0.82
FAR	0.00	0.06
Bias	0.84	0.87
ETS	0.69	0.60
KSS	0.84	0.75

$$\text{FAR} = \frac{b}{a + b}, \tag{2}$$

$$\text{Bias} = \frac{a + b}{a + c}, \tag{3}$$

$$\text{TS} = \frac{a}{a + b + c}, \tag{4}$$

$$\text{ETS} = \frac{a - e}{a + b + c - e}, \quad \text{and} \tag{5}$$

$$\text{KSS} = \frac{ad - bc}{(a + c)(b + d)}, \tag{6}$$

where a, b, c, and d are defined as in Fig. 3 and e is the correction for chance event forecasts:

$$e = \frac{(a + c)(a + c)}{a + b + c + d}. \tag{7}$$

Additional information on these skill measures can be found in Wilks (1995).

For the period of 5 April–20 September 1999, 143 days were examined using the 0–24-h (day 1) model run and 141 days were examined in the 24–48-h (day 2) run, where the difference in total days for the two forecast ranges results from missing model data (Tables 1 and 2). Rothfusz (2000) has shown that emergency managers and NWS forecasters alike deem POD (FAR) of approximately 0.80 (0.30) to be an acceptable level of performance for detection of severe convection. The POD and FAR from the MM5 convective forecasts exceed these levels at both forecast ranges (Table 2). Of particular note are the minimal false alarms at both forecast ranges (FAR of 0.00 and 0.06 for day 1 and day 2, respectively). Bias values for both periods are consistent (0.84 and 0.87 for day 1 and day 2, respectively), indicating some underforecast of observed convective occurrences. Overall, forecast performance declines minimally from day 1 to day 2, and measures of skill are high, with day-1 (day 2) ETS and KSS values of 0.69 (0.60) and 0.84 (0.75), respectively. Considering the wide range of convective cases and the total number of days sampled, these results suggest that the model forecasts can provide useful guidance to forecasters on the problem of convective occurrence.

Given the high level of skill indicated for this forecast task, it is of interest to determine the relative performance for convection that ''propagates'' into the verification domain as compared with convection that develops within that region. To examine this, we define ''propagating convection'' as an observed convective system (as defined above) that originates outside and enters into the verification domain (D3). ''Local convection'' is defined as observed convection (as defined above) that originates within D3. Of the 80 (76) observed convective occurrences at day 1 (day 2), 59 or 74% (55 or 72%) met the definition of propagating convection. Of the 13 (18) forecast misses at day 1 (day 2), however, 4 or 31% (12 or 67%) met the definition of local convection. Hence, the evidence suggests that the high skill in the prediction of convective occurrence is not governed by local circulations such as the lake breeze but may be related to the prediction of synoptic-scale factors.

Further details concerning the timing of convective occurrence will be discussed in relation to convective mode (section 3b). The remaining verifications (convective mode, section 3b; convective location, section 3c) require joint model and observed convection and represent a subset of the above dataset (67 days for day 1 and 62 days for day-2 forecasts; Table 2).

b. Convective mode

Overall, KSS for convective mode [linear, multicellular, and isolated, where the denoted mode of convection represents the most organized stage in the life cycle of each case; in an event with multiple modes, the linear mode supercedes multicellular, which in turn supercedes isolated; section 2c(2)] was 0.91 and 0.86 for day 1 and day 2, respectively. Summary measures (POD, FAR, Bias, ETS) are also reported for individual modes (Table 3), obtained by collapsing the full contingency information into 2×2 tables for the event of interest (Wilks 1995; Roebber and Gehring 2000).

Note that the difference in positive elements in Table 2 (67 at day 1 and 62 at day 2) and Table 3 (53 at day 1 and 49 at day 2) arises from the differing verification requirements for convective occurrence and mode. A positive occurrence is considered when observed and modelled convection exist for a minimum of 1 h in the forecast range. When examining convective mode, organized convection (i.e., linear or multicellular) that occurred near the beginning or end of the forecast cycle may not have persisted long enough in the model data to meet the definition of storm mode [section 2c(2)]. As a result, linear and multicellular events that did not

TABLE 3. Summary measures of MM5 forecast performance for convective mode for day 1 (0–24 h) and day 2 (24–48 h).

Forecast measure	Linear	Multicellular	Isolated
Day 1 (53 cases)			
POD	0.93	0.90	1.00
FAR	0.04	0.10	0.06
BIAS	0.96	1.00	1.07
ETS	0.80	0.78	0.91
Day 2 (49 cases)			
POD	0.96	0.77	1.00
FAR	0.07	0.18	0.00
BIAS	1.04	0.77	1.22
ETS	0.78	0.71	0.78

exist in the model range for at least 3 h were omitted from the storm-mode statistics.

The results (Table 3) indicate that, given a successful forecast of convective occurrence (Table 2), accurate detection of convective mode is obtainable on the 6-km domain (D3) for all identified modes at both forecast ranges. The prediction of the occurrence of convection, however, is itself related to the observed convective mode. Of the 13 failures to predict observed convection at day 1 (item *c* of Table 2 for day 1), 5 were isolated events (33% of the total number of observed isolated events), as compared with 7 linear (25%) and 1 multicellular (10%). At a longer forecast range, the problem of detecting isolated convection becomes more apparent. Of the 14 failures to predict convection at day 2 (item *c* of Table 2 for day 2), 8 were isolated events (89%), as compared with 5 linear (19%) and 1 multicellular (8%). Further, the four false predictions of convective occurrence (item *b* of Table 2 for day 2) were each isolated events in the model forecast. It appears that the discrimination of convective occurrence on days on which the dynamics are not distinct is subtle and becomes increasingly difficult with forecast range. This problem may be linked directly with the model physics or may represent the interaction of the model physics with initial-condition errors.

An analysis of timing errors in forecasts of convective initiation and dissipation, stratified according to convective mode, was also performed (Table 4). The results indicate a median lag in forecasts of convective initiation at both ranges. Although model ''spinup'' issues may play a role in some lag in forecasts at day 1, the persistence of this effect into day 2 suggests broader issues related to model physics, interactions between model physics and initial conditions, or other model errors. For example, Warner and Hsu (2000) showed that the use of the Kain–Fristch convective parameterization in the outer domains can limit the amount of explicit precipitation produced in the inner nest. The lag in convective initiation does not appear for isolated events; however, the sample size for this type of event is relatively small (15 and 9 isolated events at day 1 and day 2, respectively, as compared with 35 and 36 organized events for those same times), and so this statistic may be less robust. Median timing errors on the dissipation of convective events are generally within 1 h. Taken within the context of the relatively small errors in forecasts of initiation, from an operational standpoint, this performance would give a forecaster a reasonable level of confidence concerning event duration within the 6-km domain (D3). Median timing errors, stratified according to propagating versus local convection (section 3a), are similar to the overall results of Table 4, suggesting little difference in this error characteristic as a function of origination (not shown).

TABLE 4. Summary of model timing errors for convective initiation and dissipation (percentile, h) for linear, multicellular, and isolated convection. Positive (negative) values indicate model forecast convection later (earlier) than observed.

MM5 forecast	Cases	25th percentile	50th percentile	75th percentile
Linear				
0–24-h initiation	26	+5.0	+2.0	−1.0
24–48-h initiation	26	+3.0	+1.0	−4.0
0–24-h dissipation	26	+3.0	0.0	−5.5
24–48-h dissipation	26	+7.5	−0.5	−4.0
Multicellular				
0–24-h initiation	9	+2.0	+2.0	0.0
24–48-h initiation	10	+4.0	+2.0	0.0
0–24-h dissipation	9	+2.0	−1.0	−4.5
24–48-h dissipation	10	+4.0	+1.0	0.0
Isolated				
0–24-h initiation	15	+2.0	0.0	−1.0
24–48-h initiation	9	+3.0	−0.5	−3.5
0–24-h dissipation	15	+3.0	0.0	−2.0
24–48-h dissipation	9	+1.5	−2.5	−4.0

c. Convective location

The distribution of warm-season precipitation can be very erratic and often depends upon specific synoptic and mesoscale features. For events with strong dynamics (i.e., situations in which there is a substantial synoptic-scale influence on the convective environment, characterized by traditional dynamics such as lift and shear profiles, but also moisture and instability), experience suggests that the operational models (e.g., Eta and AVN) tend to give an adequate depiction of the synoptic-scale situation. Hence, there is some expectation that useful guidance can be obtained concerning the spatial distribution of convective precipitation under conditions in which strong dynamics prevail for a range of model resolutions (e.g., Anthes 1986).

The forecast becomes more problematic where larger-scale organization is lacking. For example, Lynn et al. (2001) showed that in a warm-season, widespread, moist convective case (in which heterogeneous surface forcing dominated over dynamic processes), realistic soil moisture and temperature initial conditions in addition to

TABLE 5. Threat scores (TS) and shifted threat scores (STS; see text for details) for the day-1 accumulated precipitation ≥1.25 mm for the period of 5 Apr–20 Sep 1999 where convective precipitation occurred in both the observations and the MM5 model. Quartile results are given for all convective cases and for linear, multicellular, and isolated convection.

Percentile	TS	STS
All convective cases (53 days)		
25th	0.06	0.12
50th	0.14	0.22
75th	0.34	0.46
Linear (26 days)		
25th	0.09	0.19
50th	0.23	0.31
75th	0.40	0.49
Multicellular (9 days)		
25th	0.12	0.15
50th	0.24	0.30
75th	0.55	0.62
Isolated (15 days)		
25th	0.01	0.06
50th	0.05	0.12
75th	0.09	0.15

TABLE 6. Same as Table 5 but for day 2.

Percentile	TS	STS
All convective cases (49 days)		
25th	0.06	0.15
50th	0.24	0.33
75th	0.46	0.57
Linear (26 days)		
25th	0.14	0.24
50th	0.30	0.44
75th	0.52	0.59
Multicellular (10 days)		
25th	0.05	0.09
50th	0.25	0.36
75th	0.51	0.57
Isolated (9 days)		
25th	0.02	0.02
50th	0.07	0.10
75th	0.12	0.17

sophisticated boundary layer and land surface physics were required to simulate the observed event. Thus, model resolution may be of little avail in these instances. Hence, it is of interest to assess the ''operational'' capabilities of a high-resolution model for a wide range of convective organizations.

The areal coverage of 24-h accumulated precipitation over the entire 6-km domain (covering an area of 200 448 km^2) was verified using a traditional gridpoint comparison (e.g., Olson et al. 1995). The observed precipitation was found by summing the total hourly precipitation (provided from the WSR-88D and rain gauges through the NPA) at each grid point during the corresponding 24-h period and interpolating these data (4-km grid) to the MM5 domain (D3) using a Cressman (1959) procedure. The 24-h verification allows an assessment of the aggregate effects of the convection rather than individual convective elements, a realistic goal for forecasts extending beyond the nowcast range. The effect of missing data at the four WSR-88D sites covering the 6-km model domain [Chicago, Illinois (KLOT); Milwaukee, Wisconsin (KMKX); Grand Rapids, Michigan (KGRR); and Green Bay, Wisconsin (KGRB)], which results in the gridded data being derived solely from the relatively sparse rain gauge data network, was also monitored during the course of the analysis. Missing data had minimal effect, however, because on only one study date was there downtime at more than a single WSR-88D site.

The forecast verification is based upon a threshold-based canonical threat score [eq. 4)], which is a measure of the fractional overlap between the observed and model areas meeting or exceeding a specified precipitation threshold. In this context, a is the number of grid points at which both the observed and model-accumulated precipitation meets or exceeds the precipitation threshold (T), b is the number of grid points at which the observed precipitation $< T$ but the model precipitation $\geq T$, and c is the number of grid points at which the observed precipitation $\geq T$ but the model precipitation $< T$. Because the data have been previously restricted to include only days on which convection occurred, T was set to 1.25 mm. It cannot be assumed, however, that stratiform precipitation is excluded in this analysis, because a region of stratiform rain typically occurs in conjunction with convection. Rather, the intent is to focus on the spatial distribution of precipitation on days that were dominated by convection.

Shifted threat scores (STS) were also computed. This measure seeks to account for the effect of position errors in a high-resolution domain (e.g., Roebber and Eise 2001) by conducting a stepwise shifting of the simulation domain relative to the analysis grid and recomputing the TS until a maximum value is obtained. The position shift was restricted to less than 85 km in all cases.

The TS and STS for day 1 forecasts for all convective events (Table 5) reveal performance levels that are comparable to those reported in the literature for the NCEP operational models for warm-season precipitation [e.g., Olson et al. (1995) and current Web site postings at http://www.hpc.ncep.noaa.gov/html/hpcverif.shtml]. Performance improves markedly when only the more organized modes are considered, however, with median STS increasing for linear (multicellular) modes by 41% (27%) from 0.22 to 0.31 (0.30). Stratification of the organized modes into propagating and local convection (see section 3a) reveals a slight loss of performance for the latter types, with TS of 0.19.

The results for day 2 (Table 6) show a substantial increase in skill relative to day 1, attaining median TS (STS) increases of 71% (50%) for all convective cases.

The skills are again highest for the most organized convective modes, with only limited skill for isolated events (median STS of 0.44, 0.36, and 0.10 for linear, multicellular, and isolated modes, respectively). Local convection is associated with a substantial loss of performance, with TS of 0.12. These results, for both day 1 and day 2, suggest that the in situ forcing of Lake Michigan is not elevating the model performance and that similar results should be obtainable in areas outside of the Great Lakes region.

To assess better the counterintuitive result that convective forecast skill increases with forecast range, a comparison sample was constructed that consists of those convective days for which convection was forecast at both forecast ranges and for which mode was determined (requiring a minimum of 3 h of forecast data). The comparison sample includes 46 dates, with a total of 26 linear, 15 multicellular, and 5 isolated convective events. This distribution is a representative sample of expected convective events in the warm season in this region (e.g., Changnon 1980). Median TS (STS) for all convective modes are 0.21 (0.30) for day 1 and 0.24 (0.33) for day 2. The most likely explanation for this increase in convective skill with forecast range is the removal of the deleterious effects of model spinup. Because the MM5 forecasts are initialized in a cold-start configuration (i.e., lacking condensate and, without data assimilation in the MM5 configuration, significant ageostrophic motions), it may take several hours for precipitation to develop (e.g., Colle et al. 1999). Such an effect would lead to reduced areal coverage of precipitation. One means of detecting such an effect is through the calculation of a precipitation Bias score, defined by the ratio of the number of forecast to observed grid points with 24-h accumulated precipitation amounts greater than or equal to 1.25 mm. For the 46 dates in the comparison sample, median Bias is 1.05 and 1.18 at day 1 and day 2, respectively. For those cases in which the TS at day 2 exceeds the TS at day 1 (21dates), the median Bias is 0.77 (1.13) at day 1 (day 2), lending further support to the idea that model spinup is reducing the shorter-range forecast skill.

4. Summary

The convective forecast performance of a real-time version of MM5, using 6-km grid spacing over the Lake Michigan region, was evaluated during the period of 5 April–20 September 1999. Results from this study show several positive attributes of this high-resolution data for convective forecasts:

- forecast skill for convective occurrence is high for both day-1 and day-2 forecast periods [ETS and KSS of 0.69 (0.60) and 0.84 (0.75), respectively, at day 1 (day 2)];
- forecast skill for convective mode is high for both day-1 and day-2 forecast periods, with KSS of 0.91 and 0.86, respectively [ETS of 0.80 (0.78), 0.78 (0.71), and 0.91 (0.78) for linear, multicellular, and isolated modes, respectively, at day 1 (day 2)];
- median timing errors for convective initiation/dissipation are within 2.5 h for all modes of convection at both forecast ranges; and
- the in situ forcing of Lake Michigan does not elevate model performance for convective occurrence, timing, or areal coverage, indicating that these results should be obtainable in areas outside of the Great Lakes region.

Current limitations to convective forecasting using the high-resolution MM5 data include the following:

- model spinup in the day-1 forecast period leads to reduced skill in forecasts of the areal coverage of convective precipitation at that time range, and
- the forecasts of the areal coverage of convective precipitation exhibited accuracies comparable to that of the lower-resolution, operational models (when the precipitation shield is spatially shifted to account for geographic dislocation, median TS increase to as high as 0.44 for the most organized convective modes).

The study results indicate that credible forecast information concerning convective occurrence, timing, and mode can be obtained from such models without the addition of mesoscale initial data to at least day-2 forecast range. Specificity concerning the areal coverage of precipitation remains problematic and becomes increasingly challenging as the level of convective organization diminishes. The constraints of initial-condition uncertainty and model physics errors will continue to limit explicit forecast capabilities in this arena. Ensemble approaches may help to establish the uncertainty associated with explicit model forecasts and should be fully explored.

The qualitative information obtainable from explicit forecasts, particularly in conjunction with ensemble information, may prove to be particularly valuable to forecasters. In a recent evaluation of the UWM real-time MM5 forecasts conducted by the NWS forecast offices at Green Bay and Milwaukee, forecasters found that these data allowed anticipation rather than reaction to the short-term forecast problems of convective occurrence, timing, and location. As discussed by Roebber et al. (2002) and references therein, such anticipation can be achieved through the conduct of scientific forecasting, in which a hypothesis concerning the problem of the day is established and tested prior to the issuance of a forecast. The ability to incorporate conceptual information into this process is a means by which imperfect numerical model data can be used to some advantage.

Acknowledgments. We thank Gene Brusky of NWS Green Bay for making hard copies of WSR-88D imagery available for this study. This work was supported

by the National Science Foundation (NSF) under Grant OCE-9726679 and by the University Corporation for Atmospheric Research (UCAR) under COMET Grant SOO-18116.

REFERENCES

Anthes, R. A., 1986: The general question of predictability. *Mesoscale Meteorology and Forecasting,* P. S. Ray, Ed., Amer. Meteor. Soc., 636–656.

Bélair, S., and J. Mailhot, 2001: Impact of horizontal resolution on the numerical simulation of a midlatitude squall line: Implicit versus explicit condensation. *Mon. Wea. Rev*., **129,** 2362–2376.

Bernardet, L. R., and W. R. Cotton, 1998: Multiscale evolution of a derecho-producing mesoscale convective system. *Mon. Wea. Rev.,* **126,** 2991–3015.

Blackadar, A. K., 1979: High resolution models of the planetary boundary layer. *Advances in Environmental Science and Engineering,* J. Pfafflin and E. Ziegler, Eds., Vol. 1, No. 1, Gordon and Breach, 50–85.

Bluestein, H. B., 1993: *Observations and Theory of Weather Systems.* Vol. II. *Synoptic–Dynamic Meteorology in Midlatitudes,* Oxford University Press, 594 pp.

——, E. W. McCaul Jr., G. P. Byrd, and G. R. Woodall, 1988: Mobile sounding observations of a tornadic storm near the dryline: The Canadian, Texas storm of 7 May 1986. *Mon. Wea. Rev.,* **116,** 1790–1804.

Brooks, H. E., C. A. Doswell III, and R. A. Maddox, 1992: On the use of mesoscale and cloud-scale models in operational forecasting. *Wea. Forecasting,* **7,** 120–132.

Changnon, S. A., 1980: Evidence of urban and lake influence on precipitation in the Chicago area. *J. Appl. Meteor.,* **19,** 1137–1159.

Colle, B. A., K. J. Westrick, and C. F. Mass, 1999: Evaluation of MM5 and Eta-10 precipitation forecasts over the Pacific Northwest during the cool season. *Wea. Forecasting,* **14,** 137–154.

Cressman, G., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374.

Du, J., S. L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.,* **125,** 2427–2459.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

——,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.

——,1996: A multi-layer soil temperature model for MM5. Preprints, *Sixth Annual PSU/NCAR Mesoscale Model Users' Workshop,* Boulder, CO, National Center for Atmospheric Research, 49–50.

Emanuel, K. A., and D. J. Raymond, Eds.,1993: *The Representation of Cumulus Convection in Numerical Models. Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 246 pp.

Finley, C. A., W. R. Cotton, and R. A. Pielke Sr., 2001: Numerical simulation of tornadogenesis in a high-precipitation supercell. Part I: Storm evolution and transition into a bow echo. *J. Atmos. Sci.,* **58,** 1597–1629.

Fritsch, J. M., and Coauthors, 1998: Quantitative precipitation forecasting: Report of the Eighth Prospectus Development Team, U.S. Weather Research Program. *Bull. Amer. Meteor. Soc.,* **79,** 285–298.

Gall, R., and M. Shapiro, 2000: The influence of Carl-Gustaf Rossby on mesoscale weather prediction and an outlook for the future. *Bull. Amer. Meteor. Soc.,* **81,** 1507–1523.

Geerts, B., 1998: Mesoscale convective systems in the southeast United States during 1994–95: A survey. *Wea. Forecasting,* **13,** 860–869.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 138 pp. [Available from NCAR, P.O. Box 3000, Boulder, CO 80307-3000.]

Hane, C. E., 1986: Extratropical squall lines and rainbands. *Mesoscale Meteorology and Forecasting,* P. S. Ray., Ed., Amer Meteor Soc., 359–389.

Houze, R. A., 1993: *Cloud Dynamics.* Academic Press, 573 pp.

——, S. A. Rutledge, M. I. Biggerstaff, and B. F. Smull, 1989: Interpretation of Doppler weather radar displays of mid-latitude mesoscale convective systems. *Bull. Amer. Meteor. Soc.,* **70,** 608–619.

Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 165–170.

Liu, C., M. W. Moncrieff, and W. W. Grabowski, 2001: Explicit and parameterized realizations of convective cloud systems in TOGA COARE. *Mon. Wea. Rev.,* **129,** 1689–1703.

Liu, Y., D.-L. Zhang, and M. K. Yau, 1997: A multiscale numerical study of Hurricane Andrew (1992). Part I: Explicit simulation and verification. *Mon. Wea. Rev.,* **125,** 3073–3093.

Loehrer, S. M., and R. H. Johnson, 1995: Surface pressure and precipitation life cycle characteristics of PRE-STORM mesoscale convective systems. *Mon. Wea. Rev.,* **123,** 600–621.

Lynn, B. H., and Coauthors, 2001: Improved simulation of Florida summer convection using the PLACE land model and a 1.5-order turbulence parameterization coupled to the Penn State–NCAR Mesoscale Model. *Mon. Wea. Rev.,* **129,** 1441–1461.

Nachamkin, J. E., and W. R. Cotton, 2000: Interactions between a developing mesoscale convective system and its environment. Part II: Numerical simulation. *Mon. Wea. Rev.,* **128,** 1225–1244.

Olson, D. A., N. W. Junker, and B. Korty, 1995: Evaluation of 33 years of quantitative precipitation forecasting at NMC. *Wea. Forecasting,* **10,** 498–511.

Oravec, R. J., and R. H. Grumm, 1993: The prediction of rapidly deepening cyclones by NMC's Nested Grid Model: Winter 1989–autumn 1991. *Wea. Forecasting,* **8,** 248–270.

Parker, M. D., and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale convective systems. *Mon. Wea. Rev.,* **128,** 3413–3436.

Petersen, R. A., and J. D. Stackpole, 1989: Overview of the NMC production suite. *Wea. Forecasting,* **4,** 313–322.

Reisner, J., R. M. Rasmussen, and R. T. Bruintjes, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. Roy. Meteor. Soc.,* **124B,** 1071–1108.

Roebber, P. J., and L. F. Bosart, 1998: The sensitivity of precipitation to circulation details. Part I: An analysis of regional analogs. *Mon. Wea. Rev.,* **126,** 437–455.

——, and M. G. Gehring, 2000: Real-time prediction of the lake breeze on the western shore of Lake Michigan. *Wea. Forecasting,* **15,** 298–312.

——, and J. Eise, 2001: The 21 June flood: Storm-scale simulations and implications for operational forecasting. *Wea. Forecasting,* **16,** 197–218.

——, and G. W. Reuter, 2002: The sensitivity of precipitation to circulation details. Part II: Mesoscale modeling. *Mon. Wea. Rev.,* **130,** 3–23.

——, D. M. Schultz, and R. Romero, 2002: Synoptic regulation of the 3 May 1999 tornado outbreak. *Wea. Forecasting,* **17,** 399–429.

Romero, R., C. A. Doswell, and R. Riosalido, 2001: Observations and fine-grid simulations of a convective outbreak in northeastern Spain: Importance of diurnal forcing and convective cold pools. *Mon. Wea. Rev.,* **129,** 2157–2182.

Rothfusz, L. P., 2000: National Weather Service warnings: Comparing expectations with reality. Preprints, *20th Conf. on Severe Local Storms,* Orlando, FL, Amer. Meteor. Soc., 311–314.

Sanders, F., 1987: Skill of NMC operational dynamical models in

prediction of explosive cyclogenesis. *Wea. Forecasting,* **2,** 322–336.

Steiner, M., R. A. Houze Jr., and S. E. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *J. Appl. Meteor.,* **34,** 1978–2007.

Stensrud, D. J., and S. J. Weiss, 2002: Mesoscale model ensemble forecasts of the 3 May 1999 tornado outbreak. *Wea. Forecasting,* **17,** 526–543.

——, J. V. Cortinas Jr., and H. E. Brooks, 1997: Discriminating between tornadic and nontornadic thunderstorms using mesoscale model output. *Wea. Forecasting,* **12,** 613–632.

——, J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensembles of mesoscale convective systems. *Mon. Wea. Rev.,* **128,** 2077–2107.

Wang, W., and N. L. Seaman, 1997: A comparison study of convective parameterization schemes in a mesoscale model. *Mon. Wea. Rev.,* **125,** 252–278.

Warner, T. T., and H.-M. Hsu, 2000: Nested-model simulation of moist convection: The impact of coarse-grid parameterized convection on fine-grid resolved convection. *Mon. Wea. Rev.,* **128,** 2211–2231.

Weisman, M. L., and J. B. Klemp, 1986: Characteristics of isolated convective storms. *Mesoscale Meteorology and Forecasting,* P. S. Ray., Ed., Amer. Meteor. Soc., 331–358.

Wilks, D. S., 1995: *Statistical Methods in Atmospheric Sciences: An Introduction.* Academic Press, 500 pp.

Zhang, D.-L., and R. A. Anthes, 1982: A high-resolution model of the planetary boundary layer—Sensitivity tests and comparisons with SESAME-79 data. *J. Appl. Meteor.,* **21,** 1594–1609.

——, H.-R. Chang, N. L. Seaman, T. T. Warner, and J. M. Fritsch, 1986: A two-way interactive nesting procedure with variable terrain resolution. *Mon. Wea. Rev.,* **114,** 1330–1339.