STORMTIPE: A Forecasting Experiment Using a Three-Dimensional Cloud Model

HAROLD E. BROOKS AND CHARLES A. DOSWELL III

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

LOUIS J. WICKER*

National Center for Supercomputer Applications, Urbana, Illinois

29 December 1992 and 13 April 1993

ABSTRACT

An experiment using a three-dimensional cloud-scale numerical model in an operational forecasting environment was carried out in the spring of 1991. It involved meteorologists generating forecast environmental conditions associated with anticipated strong convection. Those conditions then were used to initialize the cloud model, which was run subsequently to forecast qualitative descriptions of storm type. Verification was done on both the sounding forecast and numerical model portions of the experiment. Of the 12 experiment days, the numerical model generated six good forecasts, two of which involved significant tornadic storms. More importantly, while demonstrating the potential for cloud-scale modeling in an operational environment, the experiment highlights some of the obstacles in the path of such an implementation.

1. Introduction

Numerical models with horizontal grid spacing on the order of 1 km, capable of resolving details of individual thunderstorm structure, have proven useful in improving our understanding of convection in the atmosphere. A great deal of current severe thunderstorm knowledge has come from the use of these three-dimensional nonhydrostatic "cloud" models, coupled with theoretical and observational work (e.g., Klemp and Wilhelmson 1978; Weisman and Klemp 1982, 1984; Rotunno and Klemp 1985; Brooks and Wilhelmson 1992; Davies-Jones and Brooks 1993). The models were developed by scientists within the meteorological research community, and until recently there have been no efforts to use them in an operational (or pseudooperational) setting. This is in sharp contrast to the so-called "mesoscale" models, usually hydrostatic, with horizontal grid spacing on the order of 20–50 km, which in many cases were developed by scientists working for organizations with missions to support operational meteorology. As a result, the mesoscale models have been used as what might be considered high-resolution numerical weather prediction (NWP) models in a pseudooperational mode, in which the modeler can work with the initial conditions in an effort to reproduce the observed weather. The resultant products resemble current NWP products, except at higher resolution.

As computer power increases, the distinction between mesoscale models and "traditional" larger-scale NWP models will become blurred. It appears likely, however, that there are large obstacles in the path of using models with the resolution of current research cloud models in an NWP mode. [For arguments for and against such use, see Droegemeier (1990) and Brooks et al. (1992), respectively.] This is not to say that, with increasing computer power, there will be no useful role for cloud models in an operational environment. From 21 March to 30 May 1991, research scientists and operational forecasters from the National Severe Storms Laboratory (NSSL), the National Weather Service Forecast Office in Norman, Oklahoma (NWSFO OUN), and the National Center for Supercomputer Applications (NCSA) carried out an experiment to look at one possible way in which cloud models could be used operationally. This paper is a preliminary report on the design and results of this experiment, known as the Storm Type Operational Research Model Test Including Predictability Evaluation (STORMTIPE). We describe briefly the general procedures followed in the experiment, including the development of the forecast and verification issues. We follow with a discussion of the results and specific examples of both "successful" and "unsuccessful" model forecasts. Finally, we conclude with general remarks concerning the implications for operational use of cloud-scale models.

* Current affiliation: Department of Meteorology, Texas A&M University, College Station, Texas.

Corresponding author address: Dr. Harold E. Brooks, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069.

2. Experimental procedure

a. Forecast development

On days when the experimental forecasters (a team consisting of one forecaster from NSSL and one from NWSFO OUN) anticipated strong or officially severe (defined by a wind gust of at least 25 m s^{-1}, hail of diameter 19 mm or greater, or a tornado) convection within the state of Oklahoma, excluding the Panhandle, the forecast team produced forecast environmental conditions for later in the day for a specific time and place near where and when they expected mature convection (see Fig. 1 for a map of locations mentioned within the text). These conditions included a surface pressure as well as thermodynamic and wind profiles up to approximately 100 hPa. This product was completed at about 1100 LST (1700 UTC before 7 April and 1600 UTC after), and was valid for a time typically between 2200 and 0000 UTC, although one forecast was for as early as 2000 UTC. Forecasters were allowed to use any data routinely available to operational forecasters to assist in the preparation of the forecast. Minor modifications, if necessary to satisfy the input needs of the numerical model, were made to this forecast by noon local time, and the forecast was converted to digital form by one of the authors (HEB). The most common of the modifications was the elimination of an inversion at the top of the boundary layer. The forecast sounding was then used as input for the cloud model of Wicker and Wilhelmson (1990), and a simulation of 2 h of storm development was carried out on the CRAY-2 supercomputer at NCSA, generally finishing before 1400 LST. The model is similar in formulation to the Klemp–Wilhelmson cloud model (Klemp and Wilhelmson 1978) but with more accurate numerical techniques and with a microphysical parameterization including three categories of ice, in addition to water vapor, rainwater (large liquid drops), and cloud water (small liquid drops) (Straka and Anderson 1993). The vertical grid spacing near the ground was 400 m, and the grid was stretched so that the spacing was 600 m near the top of the model. The horizontal grid spacing was 1500 m. A variety of products was generated automatically by the model and analysis software, and at the conclusion of the simulation, one of the modelers involved in the project (LJW or HEB) would interpret the model results and generate a brief model guidance report. This product summarized the model output, typically including predictions of the storm type (e.g., supercell or not, tornadic or not), storm motion, and chances for and size of hail. This subjective guidance, along with the objectively generated products, was then presented to the experimental forecast team.

b. Verification issues

There are two basic types of errors that can cause the final model guidance to be wrong. Our efforts at

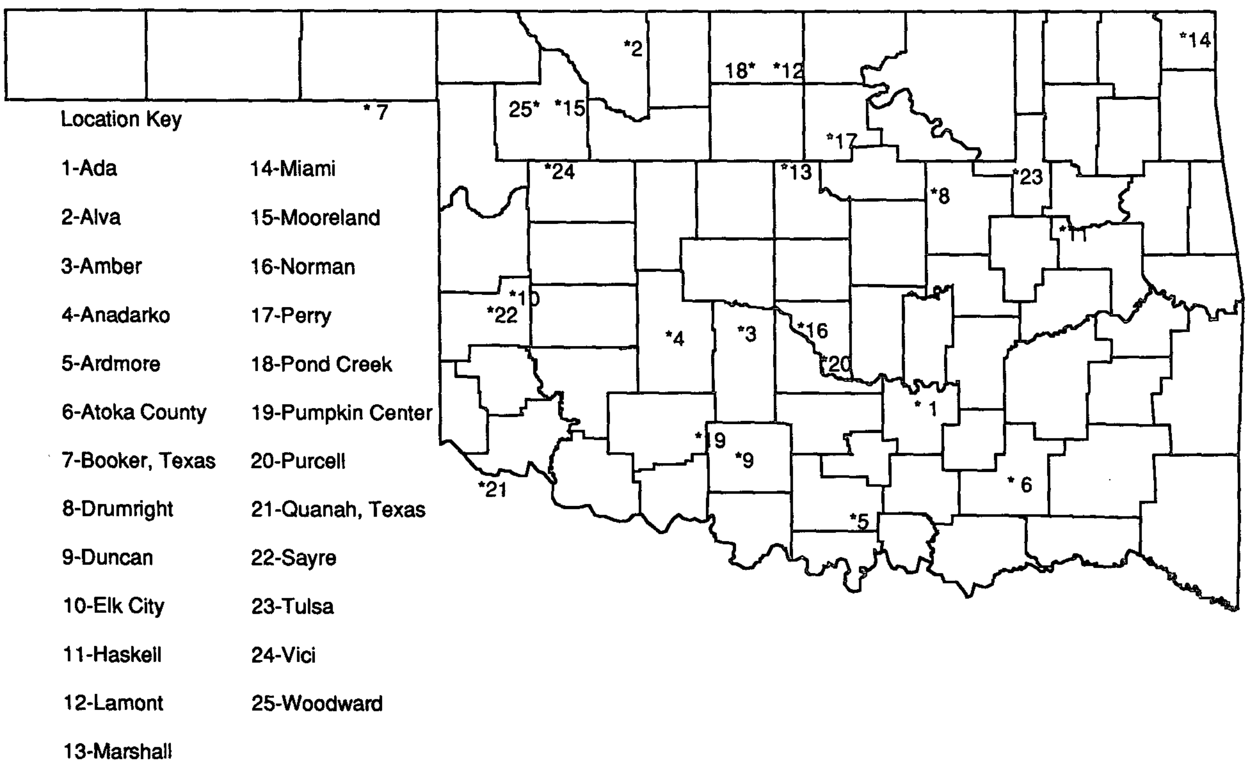

FIG. 1. Map of Oklahoma with locations mentioned in the text.

verification were designed to try to isolate the effects of these errors. The first class of error (type I) involves the generation of an inaccurate forecast of the environmental conditions. The second is that, even with an accurate forecast of the environment, the numerical model may be incapable of generating a storm that resembles the storm observed in association with the accurate forecast environment (type II). Verification was separated along these lines. For reasons that will become apparent, some of the verification was only qualitative in nature.

In an effort to quantify type I errors, several potential data sources were used to verify the sounding forecast. The first and most important was the use of the NSSL mobile ballooning facilities, the Mobile Cross-Chain LORAN (Long-Range Aid to Navigation) Atmospheric Sounding System (M-CLASS) (Rust et al. 1990). Teams were sent out to the forecast sounding location with instructions to launch serial soundings every 60–90 min from the forecast location prior to and including the forecast time. Unfortunately, this procedure did not work as well as desired. Beyond the standard problems associated with radiosonde data collection, we encountered additional difficulties. Early in the experiment, there were problems in getting LORAN data necessary for wind measurements in the vicinity of convection. Clearly, for our purposes, this was a major limitation. Later during the experiment, another NSSL field project had priority for the use of the mobile sounding vehicles. As a result, soundings were as much as 150 km from the location specified in the forecast. All in all, on most occasions we were unable to get the quality and quantity of mobile balloon soundings desired.

A second verification data source (for type I errors) was the operational National Weather Service (NWS) sounding taken at Norman. The spring of 1991 was particularly active in northern and western Oklahoma, so that, particularly on days when mesoscale variations might be large, the Norman sounding was not necessarily representative of the environment near the storm.

A third data source for wind information was the hourly averages of wind from the Profiler Demonstration Network, from which the data became available as the project was under way. Four sites in Oklahoma (Vici in the northwest, Lamont in the north, Haskell in the east, and Purcell in the central part of the state) were relevant to our experimental domain, and provided data at various times during the project. Owing to problems with the network, data were not always available. When available, they were extremely useful in detailing the vertical wind profile, a particularly important parameter in determining the type and organization of convection in the atmosphere.

Verification data used to assess type II errors was more subjective and qualitative than that for type I errors. One of the strong points of restricting the forecast region to Oklahoma is that the NWSFO OUN is aggressive in their pursuit of severe weather reports. As a result, it is unlikely that many significant events could have gone undetected in the forecast region. These reports gathered by the NWS formed the basis of our verification for severe weather days. Doppler radar data from the WSR-88D radars also were available to give indications of significant midlevel rotation and the degree of organization of storms on many of the forecast days. On nonsevere weather days, any evaluation of the nature of convection depended on field observation by storm chasers. We have classified the type of storm subjectively, based on all of these data. It was possible, in some cases, to attempt to quantify parameters such as the storm motion. Uncertainties in the radar observations made this troublesome occasionally, and more analysis into quantitative measures of type II errors is needed.

One of the more interesting possibilities was that errors of both kinds could occur and still result in a qualitatively good forecast. For instance, if the forecast sounding had overestimated the amount of moisture available and the numerical model took that moist sounding and made a verified forecast of supercell convection, while with the observed moisture profile it failed to produce convection, the combination of the two errors might lead to a good forecast. This problem becomes important in considering those days where we have no, or poor, verification soundings. It is tempting to assume that an accurate model forecast could come only from an accurate input sounding. Unfortunately, we can have no such assurance. The behavior of the numerical model and its sensitivity to changes in the input conditions are extremely complex and not well understood at this time. It may be true that a wide variety of bad initial conditions could produce qualitatively accurate results, while qualitatively good initial conditions, with only small errors, could produce poor forecasts. Thus, we recognize that our attempts to "verify" the forecast sounding by noting that the model produced a qualitatively good forecast is not rigorous. However, for cases without sounding data, that was all that could be done, and we want to make the caveats clear.

3. Results

a. General

Several members of the sounding forecast team had little experience in attempting to make pinpoint forecasts of sounding and hodograph data. As a result, particularly in the early stages of the experiment, forecasters occasionally struggled with the procedure and ended up with little time in which to attempt to make a forecast. Similarly, the numerical modelers were inexperienced at first in the production of model simulations at such a rapid rate. Thus, the process of conversion of the sounding forecast into a model-usable

TABLE 1. Description of days when forecast sounding was made. Asterisks indicate days on which no model forecast was run. The *G* indicates good verification soundings. Locations are all in Oklahoma unless otherwise noted. The *U* indicates that the numerical model underforecast the severity of the event, and *O* indicates it overforecast the severity. The *X* indicates a qualitatively good forecast.

Date		Req. location	Error type	Comments
21 March		Central Oklahoma	I	Ada, Atoka Co. (southern Oklahoma) tornadoes
24 March	*	Norman		No convection
2 April	*	Ardmore		Duncan, Pumpkin Center weak tornadoes
8 April	*G*	East of Tulsa	*X*	Hailstorm near Miami
11 April	*G*	Sayre	£	No long-lived convection
12 April	*G*	Western Oklahoma	II (*U*)	Pond Creek, Marshall tornadoes
17 April	*G*	Anadarko	*X*	Severe, short-lived cells
24 April		Southwestern Oklahoma	II (*O*)	MCS, tornado near Quanah, Texas
26 April		Amber	I	Violent tornado outbreak from Oklahoma to Nebraska
28 April		Norman	#	Small hail in eastern Oklahoma
3 May	*	Drumright		Late sounding forecast
4 May		50 miles west of Oklahoma City	I	Nontornadic supercell in southern Oklahoma
15 May	*G*	Elk City	*X*	Shamrock, Texas, and Laverne tornadoes
16 May		Perry	II (*U*)	Wichita, Kansas, and Tulsa tornadoes
26 May		Alva	*X*	Woodward–Mooreland tornado
29 May	*	Woodward		Booker, Texas, weak tornado
30 May	*	Woodward		No convection in Oklahoma

On 28 April, an "optimistic" sounding produced a forecast convective line with a supercell on the end. The supercell never developed, although a line of marginally severe thunderstorms did.

£ Two forecasts were run, with the first producing a good forecast of the convection and the second significantly overforecasting the severity of the weather.

format improved as the experiment went along, as did the interpretation and transmission of model results.

Sounding forecasts were generated on 17 days (Table 1). We could have confidence in the M-CLASS verification of those soundings on only 5 days (8 April, 11 April, 12 April, 17 April, and 15 May). On the others, the mobile soundings were either contaminated by convection, of poor quality due to data acquisition problems, or clearly in the wrong location—for example, 200 km from the forecast location or on the dry side of the dryline when the forecast was for the moist side. We have considered the most unstable observed sounding on the 5 days for which we have good verification, and considered the forecast sounding at selected heights. In general, the forecasts were too warm and moist at low levels and too cold at midtropospheric levels, resulting in an overforecast of the instability. The forecast wind profiles, on average, were backed more than observed, resulting in an overforecast of the environmental helicity (Fig. 2 and Table 2). As a result of the overforecasts of instability and helicity, the forecast soundings tended to be more favorable to supercell convection than the M-CLASS observed soundings. It is interesting to note, in this context, that the combination of shear and instability in the form of the bulk Richardson number (BRN) indicates that none of the forecast environments are associated with the range of BRN typically associated with supercells (Weisman and Klemp 1982, 1984). This overforecasting tendency results, at least in part, from the tendency of severe storm forecasters to predict the "worst case" forecast. This is not entirely without value in real world forecasting, where the penalty for failing to anticipate the worst that can happen is greater than that for overforecasting (see Murphy 1991). The experimental forecasters were not discouraged from engaging in this practice, since we hoped to run the model forecasts on *all* potentially severe thunderstorm days. Therefore, this overforecasting bias was not unexpected.

Of the 17 forecast days, numerical model forecasts were run on 12. No forecast was made on a day when there was no one available to run the model or on a

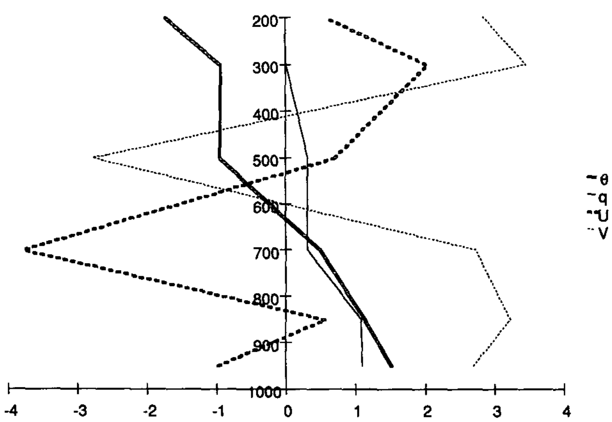

FIG. 2. Average errors for forecast soundings verified by M-CLASS balloons for 8, 11, 12, and 17 April, and 15 May 1991. For potential temperature (θ) horizontal scale is in Kelvin; for water vapor (q), in grams per kilogram; and for horizontal winds (u and v), in meters per second. Errors are calculated at the surface, 850, 700, 500, 300, and 200 hPa.

TABLE 2. Parameters of forecast (Fcst) and observed (Obs) soundings for days with good verification soundings and for 26 May 1991 case. BRN is bulk Richardson number (Weisman and Klemp 1982, 1984), and helicity is storm-relative 0–3-km helicity (Davies-Jones et al. 1990). Wind parameters on 12 April and 26 May taken from nearest profiler site.

	CAPE		Lifted index		BRN		Helicity	
Date	Fcst	Obs	Fcst	Obs	Fcst	Obs	Fcst	Obs
8 April	3248	2299	−9.2	−6.1	52	23	199	79
11 April	3988	2409	−8.3	−7.5	69	132	179	142
12 April	3589	2182	−10.0	−6.5	62	39	210	243
17 April	4474	2623	−9.9	−7.6	104	104	129	73
15 May	5080	2625	−10.9	−7.4	69	23	242	315
26 May	3897				92		194	155
Average	4046	2428	−9.4	−7.0			192	168

day when the forecast sounding was not available before 1400 LT. On 1 day early in the experiment, the forecast team chose to produce a forecast sounding and to request verification, even though they expected no significant convection, in order to practice procedures. Finally, no model forecasts were made on the last 2 days of the experiment, because of limitations on computer resources for the project. A total of 13 model forecasts were made, with two run on 11 April, as will be discussed later. Three of the initial environmental forecasts clearly were poor (type I error), resulting in unrealistic model forecasts. In one, the return of low-level moisture was underestimated; in another, the intensification of the low-level wind profile curvature was underestimated; in the third case, cloud cover from earlier convection apparently prevented surface temperatures from rising to the forecast level. Because of the previously discussed problems with gathering verification data for the soundings, we cannot be certain that all the other forecast soundings were good. That may have some impact on the estimation of type II errors from the numerical model.

The numerical model overforecast the severity of the event in two cases, underforecast in two others, and made qualitatively good forecasts in six other cases, two of which were tornadic forecasts. Of particular importance to future work are the events of 11–12 April. On 11 April, the sounding from the forecast team, containing a relatively high level of free convection (LFC), was run in the model. The initial temperature impulse used to start the storm rose to the equilibrium level of the sounding, producing a strong updraft ($\sim$30 m s^{-1}) that failed to sustain itself for more than an hour. [See Brooks (1992) for a discussion of the problems numerical thunderstorm models have with high-LFC environments.] The modeling team decided that that forecast was "unrealistic" and increased the low-level moisture slightly, lowering the LFC. This second forecast resulted in a simulated classic supercell, with significant, long-lived low-level vorticity, indicative of a strong likelihood of tornadic activity. In reality, the actual convection in western Oklahoma resembled the initial model forecast, with rapidly growing updrafts producing anvils after a short time, but with the lower part of the storm evaporating almost as quickly, leaving "orphan anvils." (The verifying sounding indicated that the original moisture profile had been more accurate.) On the following day, the model forecast resembled the previous day's first forecast. As a result of the events of 11 April, the modelers decided not to change the sounding and rerun the model. On 12 April, however, three thunderstorms developed in central Oklahoma. The southernmost storm followed the evolution of the storms of the preceding day after briefly producing severe hail. The other two storms, however, developed into supercells that produced tornadoes in the late afternoon and evening in northern Oklahoma, which the model had not forecast as likely. We will discuss this case in more detail later.

b. *Specific cases*

1) 26 MAY 1991

To show the numerical model products, we will use as an example one of the two successful forecasts of tornadic storms, the case of 26 May 1991. It also illustrates the difficulties of the sounding verification procedures. An outflow boundary approaching from the north and a dryline approaching from the west were forecast to produce a "triple point" in northwest Oklahoma. Using this as the main feature of interest, the forecast team produced a sounding and hodograph valid at 0000 UTC that evening at Alva, Oklahoma (Fig. 3). With that sounding, the numerical model forecast generated a classic supercell with a strong, long-lived low-level mesocyclone. Supercells developed in the region late in the afternoon, with one producing 10-cm diameter hail and three tornadoes, including a strong tornado (F2-F3) between 2300 and 0000 UTC in the vicinity of Woodward and Mooreland, Oklahoma, about 75 km southwest of Alva. The model storm moved almost due east, in the same direction but slightly faster than the observed storm (10 m s^{-1} compared to 7 m s^{-1}). In most respects, this forecast was a spectacular success.

To verify the forecast conditions, we have used data

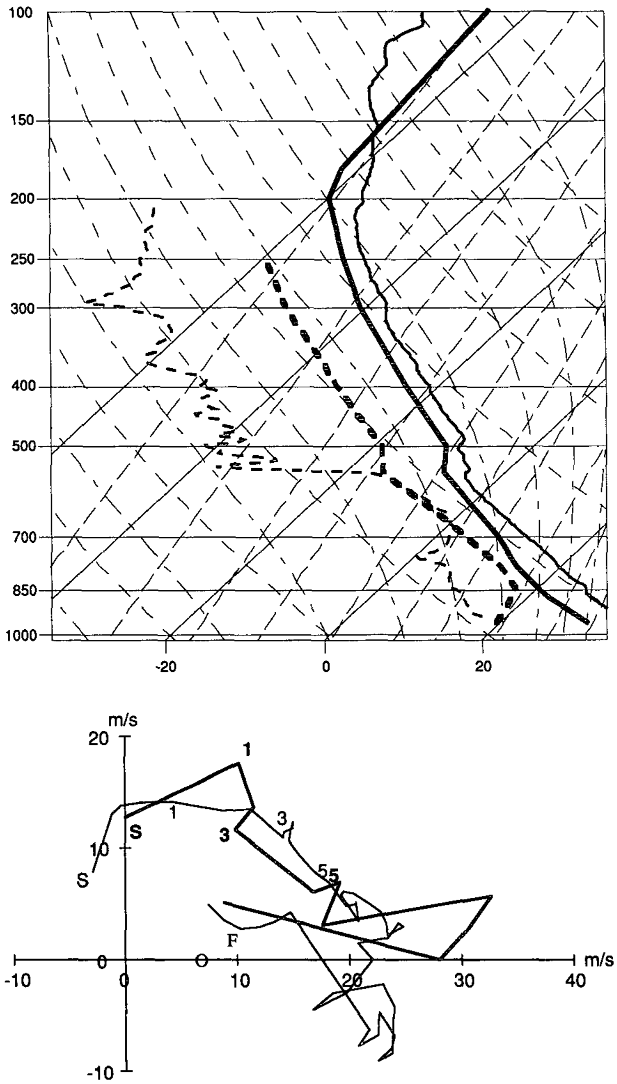

FIG. 3. Forecast (dark lines) and observed (light lines) environmental conditions for 2300 UTC 26 May 1991. (a) Skew T–log$p$ diagram with Erick sounding information (temperature—solid lines; dewpoint—dashed lines) for observed thermodynamics. (b) Hodograph with Vici profiler data for observed winds. The S indicates surface winds, and numbers indicate approximate heights above ground in kilometer. The F is model-generated storm motion, and O is approximate observed storm motion. Scale is in meters per second.

from the wind profiler network station at Vici (approximately 40 km south-southeast of Woodward) and the M-CLASS sounding at Erick, Oklahoma, (nearly 150 km south-southwest of Woodward, being used by another experiment on the day) from 2300 UTC. We are using the profiler data not only because it is closer to the event of interest but also because the mobile sounding failed to get good wind measurements as a result of problems with determining the position of the balloon using Loran, although the thermodynamic data were good.

While the forecast wind profile is relatively accurate over the lowest 5 km, the thermodynamic structure, particularly the moisture profile, is not close to the observed. In this case, we have reason to believe that the measured sounding is unrepresentative of the storm's environment. The Woodward storm was the southernmost convection in the region, and it is possible that the low-level moisture was significantly higher in the vicinity of the storm than in the region farther south where the verification sounding was taken, increasing the instability of the environment in that region. This is a clear case in which our lack of control over the location of the verifying sounding has hindered our ability to evaluate the forecast.

In any event, we can use this case to show the output that was made available to the forecast team from the numerical model. We produced two summary products and an animation of the model field evolution. The first summary product was a brief written description of our interpretation of the model forecast (Fig. 4). Our goal was to give the forecasters the benefit of our experience with the numerical model and to "translate" its output into parameters that could be compared with observations—for example, storm motion, storm type, and associated weather. We also tried in many cases to indicate our confidence in the forecast, particularly on those days when we were suspicious of model performance or when it had done something we had not seen before. The second summary was a time history of the maximum updraft in the model domain and maximum absolute vertical vorticity at the lowest model level (Fig. 5). These two quantities were important in the development of our interpretation, and gave us indications as to whether the storms were long-lived and well organized. In this case, the sustained high value of low-level vorticity was a strong indicator of the intensity of the low-level circulation. To illustrate the nature of the forecast storm further, note the distinct hook echo and associated vorticity maximum associated with the cell in the center of the domain near the end of the simulation (Fig. 6). The structure was clearly supercellular. (The northern cell in the domain is the left mover from the splitting storm that also produced the right-moving supercell in the center of the domain. The left mover, while showing a possible inflow notch and hook echo, did not have significant

26 May numerical model forecast

Splitting storm develops (Right-mover moves at 260°/20 kts)

Significant low-level vorticity for a long time suggestive of possible tornadoes

Much hail aloft, with some making it to ground (model may melt hail too quickly on descent)

Brooks

FIG. 4. Example of worded guidance associated with 26 May STORMTIPE forecast.

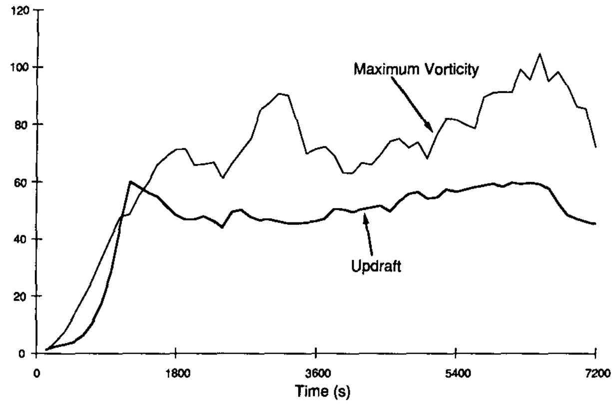

FIG. 5. Time history of maximum updraft (bold—m s^{-1}) and low-level vorticity (light—10^{-4} s^{-1}) for 26 May 1991 STORMTIPE forecast.

low-level vorticity. Its presence in the domain, however, serves to illustrate the complexity of even a relatively straightforward and simple numerical thunderstorm simulation in an operational environment and the need for model interpretation in the forecasting process.)

Animation of the evolution of the updraft and radar reflectivity[1] at 1-km altitude in the model were produced with a time step of 5 min and shown on a Macintosh II computer loaned by NCSA to the experiment; we cannot provide the animation in this paper. However, our experience indicated that this, or a similar product, provides a good interface between the operational forecaster (who is used to seeing radar loops) and the numerical model output. As Doppler radars become operational in the future, it will be useful to consider the animation of reflectivity and single-Doppler velocity as starting points for integrating storm-scale numerical model output in a forecast regimen.

2) 17 APRIL 1991

A qualitatively good forecast of the type of convection on a nontornadic day was made on 17 April 1991, a day on which the forecast sounding agreed reasonably well with the observed sounding (Fig. 7). The environment was uncapped and convection was forecast correctly to begin early in the afternoon. The model forecast showed relatively strong, short-lived convection (Fig. 8) and indicated significant amounts of hail. The observed storms developed and became severe rapidly, producing large amounts of 2-cm diameter hail and brief, strong surface winds. As each storm died 1–2 h later, its outflow produced additional storms that followed the same life cycle as their predecessor. The situation created a significant (and not uncommon) operational warning problem. Storms met severe weather criteria within perhaps 20–30 min of the first echo, but each storm was severe for only about an hour or less. The subsequent redevelopment of convection led to brief, repeated episodes of severe weather scattered over a large area, with much of southern Oklahoma under severe thunderstorm warnings during the afternoon. Because of the short time between the first observed echo from a storm and the brief duration of any of the individual events, operational forecasters were faced with the option of having either a large false alarm rate or failing to have any warnings out on some storms.

The model simulation captured essential qualitative elements of individual storms on this day, with rapid development of strong, but short-lived convection.[2] While such a model forecast could be useful in describing the *general* character of anticipated convection, the significant questions of defining convective redevelopment locations and the intensity of the redeveloping storms remain unanswered. Since the observed weather resulting from the whole family of storms is sensitive to the detailed initial storm evolution and the subsequent redevelopment, the utility of the forecast is limited. Also, because it is doubtful that we ever will know those details sufficiently well, it is likely that models using explicit prediction, even with early observations from Doppler radars, will have a difficult time in the foreseeable future producing a more useful forecast on such days than the *qualitative* forecast produced by the model. For an extended discussion of this issue, see Brooks et al. (1992).

3) 12 APRIL 1991

As mentioned already, the STORMTIPE forecast for 12 April 1991 was for little chance of long-lived convection. In reality, a supercell storm, producing multiple tornadoes over a period of a few hours, passed approximately 25–30 km to the west of the Lamont profiler and a strong-to-violent tornado (F4) was on the ground at that time. A second storm, farther south, produced two or three small, short-lived tornadoes. Although the observed surface winds are backed more than in the forecast and the low-level moisture was overforecast, the temperature profile is an excellent forecast (Fig. 9): errors at 700 hPa are on the order of 1 K. Note that, in this case, the verification sounding was taken approximately 150 km south of the most

[1] While the model predicts moisture variables and not radar reflectivity, the reflectivity is diagnosed using a Z–R relationship (Straka and Anderson 1993). The physically important quantities, as far as storm dynamics and structure are concerned, are the predicted moisture variables, but the diagnosed radar reflectivity was presented to the forecasters in an effort to provide them with a product resembling radar output, with which they were more familiar.

[2] This situation is referred to colloquially as "nuclear popcorn" convection.

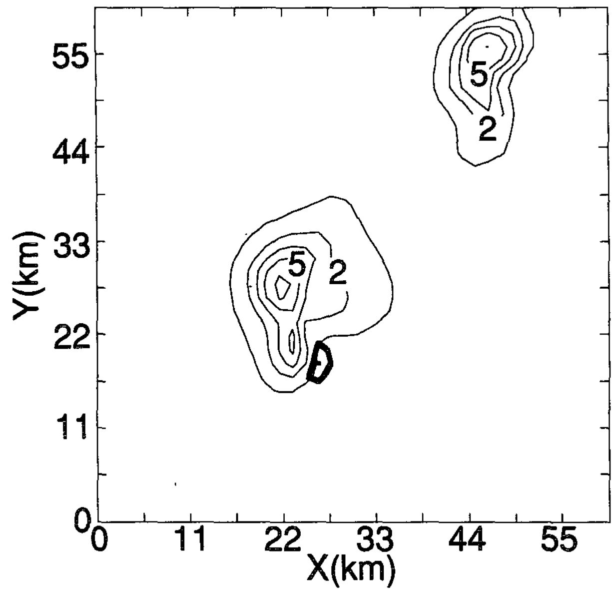

FIG. 6. Rainwater (light) and vorticity (dark) at 100 m at 6600 s of 26 May 1991 STORMTIPE forecast. Tick marks are 5.5 km apart. Contour interval of 1.5 g kg^{-1} for rainwater (lowest contour 0.5 g kg^{-1}) and 0.005 s^{-1} for vorticity with zero contour suppressed.

intense storm and east of the location of the storm that dissipated relatively quickly. As a test of the sensitivity of the model to small changes in the initial conditions, we reran the model twice in a hindcast mode, keeping the forecast moisture profile throughout the experiments, and including first the observed temperature and then both the observed wind (from the Lamont profiler, including the surface winds) and temperature. (The moisture profile was not changed since the initial model forecast had not generated a long-lived storm and it was felt that decreasing the moisture would not produce a storm either.)

When the observed temperature was used with the forecast moisture and hodograph, the model storm became long-lived and produced two relatively short-lived vorticity maxima (close to, but not quite achieving the 10^{-2} s^{-1} "mesocyclone threshold") at low levels (Fig. 10). (Compare the vorticity to the long-lived maximum in the 26 May 1991 forecast in Fig. 5.) Qualitatively, this storm was similar to the southern tornadic (weaker) storm with a weak low-level mesocyclone. When the observed profiler winds also were included with the observed temperature and forecast hodograph, the storm failed to generate significant low-level vorticity and provided no hint of tornadic potential at all.

This case is particularly interesting. The first point of interest is the extreme sensitivity to the initial temperature profile, even without considering the large differences between the forecast and observed moisture profiles. The implication is that any operational use of the current generation of cloud models would require, at least in some situations, temperature profiles accurate to 1 K or better in order to make a forecast of the occurrence or nonoccurrence of convection, let alone a forecast of the associated weather. The second relates to the wind profile information. The proximity of the wind profiler to the northern storm and its location in (presumably) the inflow region of the storm would lead one to believe that an extremely accurate depiction of the winds would be available. Given that the model fails to produce anything close to strong rotation at low levels, we are left with two options. The first is that the profiler failed to detect the environment of the inflow into the storm. If so, the problem of detecting supercell environments operationally is an extremely difficult one, because this case implies that proximity of 30 km is insufficient in some instances. The second option is

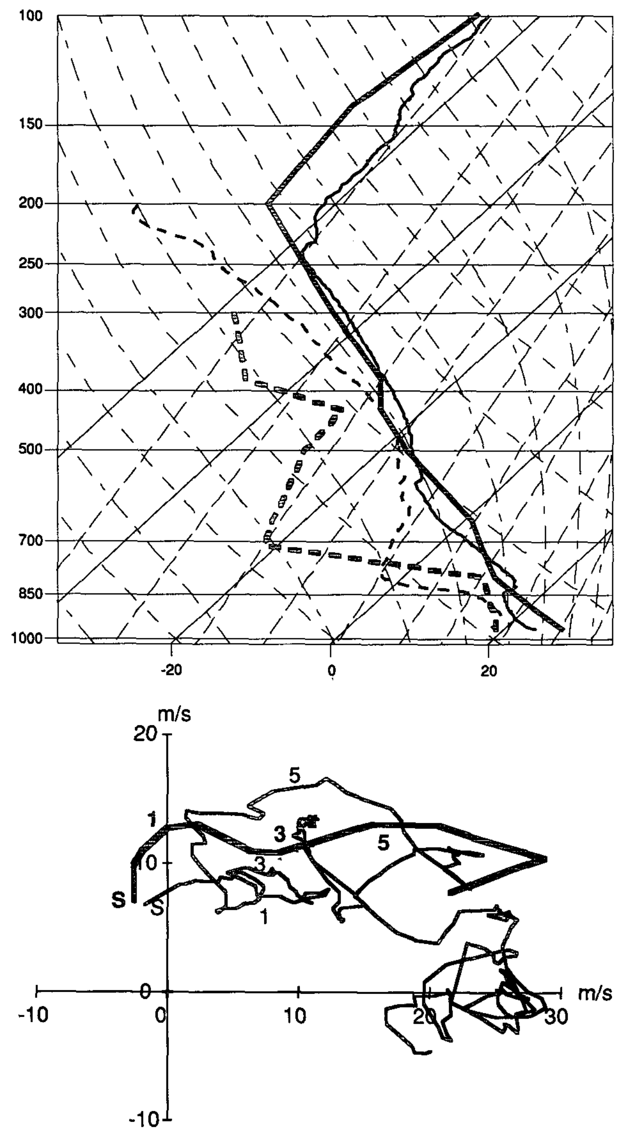

FIG. 7. As in Fig. 3 except for 1900 UTC 17 April 1991. Observed winds taken from the M-CLASS sounding.

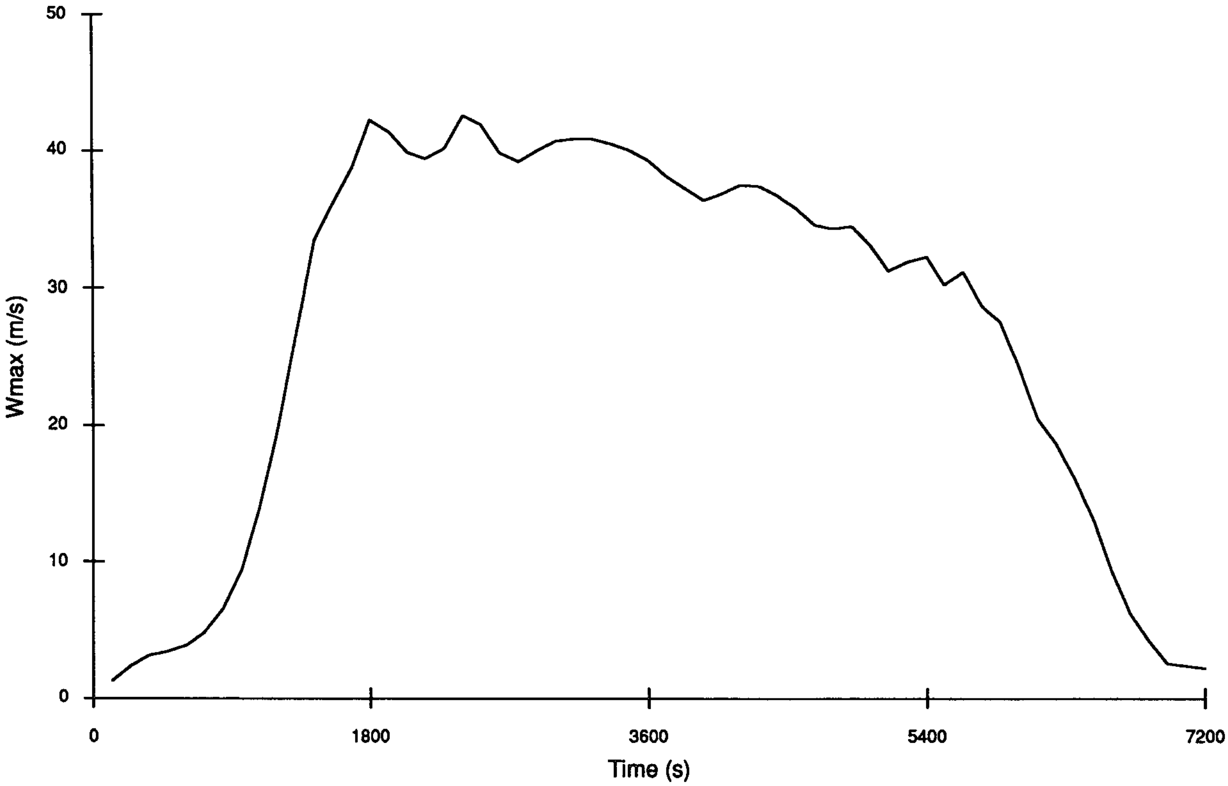

FIG. 8. Maximum updraft in 17 April 1991 model forecast.

that there are gaps in our understanding of the physical processes leading to tornadogenesis. While this notion is no more comfortable than the first one, it at least offers hope of a practical solution.

4. Discussion

STORMTIPE was a successful experiment in many ways, particularly in indicating important areas for future work, if cloud-scale numerical models are ever going to be part of operational meteorology. The numerical model forecasts appeared to be superior to some simple methods of sounding analysis, particularly in the forecasting of supercells (e.g., the cases of 15 May and 26 May). On the side of the human forecaster, the great difficulty of making detailed accurate forecasts of the environment is crucial. The 12 April storm has shown that whatever method is used to provide an initial environmental profile (i.e., human forecasters or large-scale numerical models) there will be cases in which the model may be extremely sensitive to the initial conditions. Similarly, the 17 April storms, with their short, severe life cycles, illustrate the difficulty in some environments of using even an accurate forecast. Methods to assist the forecaster, be they better training or improved large-scale numerical models or something else, need to be developed as the need for forecasting convective-scale weather increases. As an example of the problems associated with the large-scale models, the nested grid model from the National Meteorological Center did not provide good forecasts of boundary-layer moisture, one of the most important variables in determining the potential for convective development. A critical issue in the interpretation of our results is that frequently there was more than one storm in the forecast domain. Thus, our verification is subjective in the sense that we have had to choose *which* storm to verify against. In this case, we always have taken the most severe storm; in an operational setting, however, *all* of the storms may be important. The use and interpretation of the model data would then become more complicated in a nonexperimental setting.

We see several important areas in which cloud-scale models can be useful, even without attempting to use them in an explicit forecast mode, à la NWP models. Brooks et al. (1992) proposed that a more effective use of these models is in a quasi-Monte Carlo mode, with a set of model forecasts covering a reasonable range of possible environmental conditions. The 11 April simulations give a concrete example of how such a procedure might be useful. Long-lived convection occurred only when ample low-level moisture was present. Assuming that moisture was the only parameter over which model behavior varied significantly, the operational forecaster might then pay particular attention to dewpoints in the forecast region during the day. If

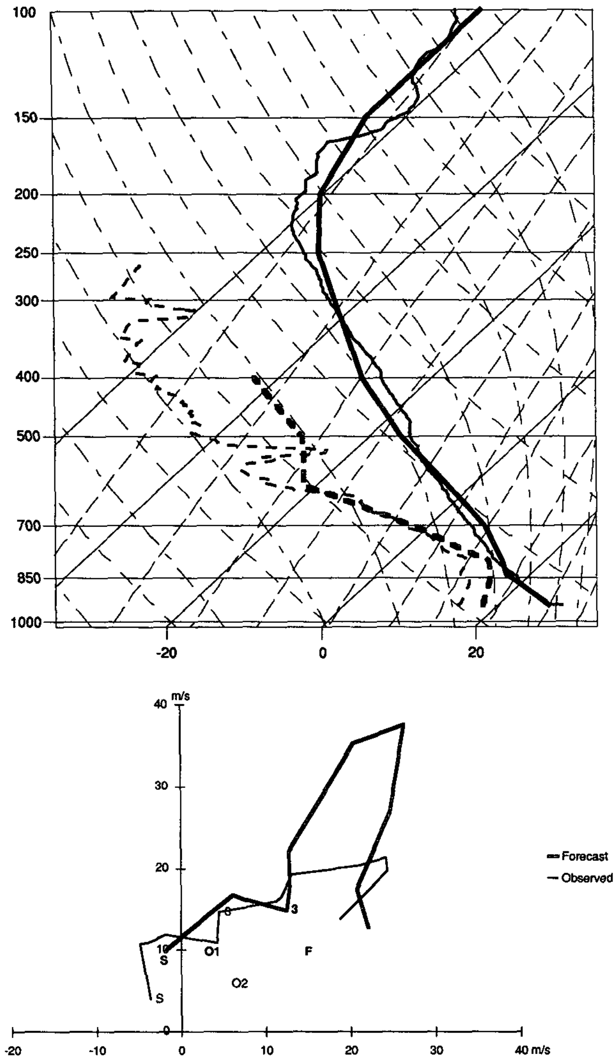

FIG. 9. As in Fig. 3 except for 2300 UTC 12 April 1991. (a) Verification sounding from southern Oklahoma. (b) Verification hodograph from Lamont profiler. Here O1 is the observed motion of the storm that produced strong tornadoes in northern Oklahoma, and O2 is the motion of the storm producing weaker tornadoes farther south. The forecast storm movement is taken from the simulation producing the longest-lived low-level vorticity maximum.

they rose to high levels, the forecaster might anticipate the development of severe convection and be able to alert weather spotter groups to provide assistance in monitoring the situation.

Closely related to this issue is the potential for the model to define probable and improbable events. This potential forces the forecaster to consider the possibility of rare events. By doing so, the forecaster and public are less likely to be surprised by the evolving weather situation.

Numerical models of convection, whether or not they are used operationally, can be used in training of forecasters. By seeing how varying a parameter affects the evolution of storms, the forecaster can learn the importance of that parameter within the context of a particular environment. Also, the model can allow a forecaster to "experience" a wider range of storm types than may be encountered in regions where severe weather events are rare. Again, this can be useful in helping forecasters anticipate (and hence recognize) events, particularly with severe weather when it can be difficult to make a good forecast but when public safety requires one.

Moreover, it is imperative that developers of potential operational small-scale models interact with operational forecasters from an early point in the process. The experience of the operational community can be important in determining what kind of products from numerical models are useful. The need for the products to be assimilated quickly and to be interpreted easily is clear. It is not always obvious to model builders that the kinds of products that are useful in an analytical research mode are not necessarily the same as those that are useful in a forecast mode. By involving operational forecasters in the testing stage of models, the utility of the model may be maximized.

We plan to expand our efforts in this direction in

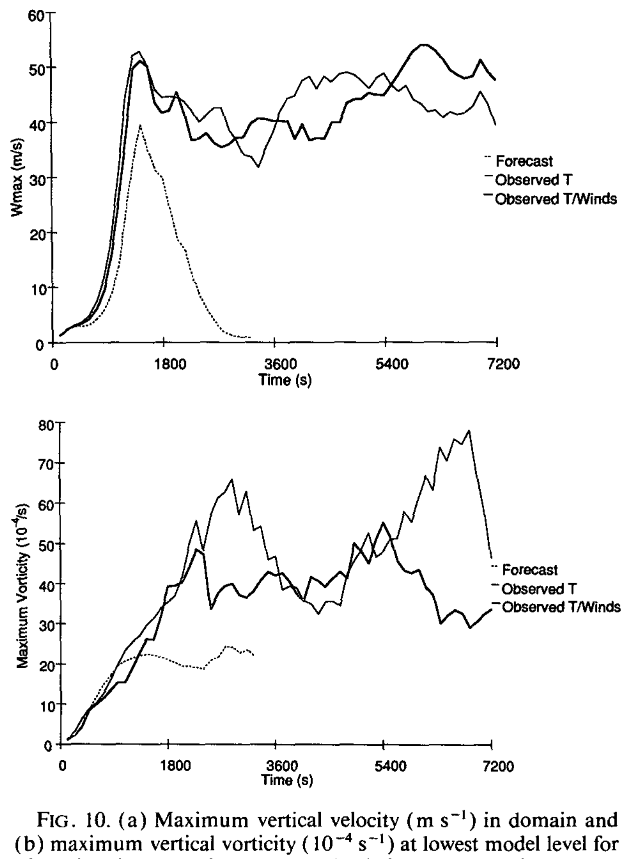

FIG. 10. (a) Maximum vertical velocity (m s^{-1}) in domain and (b) maximum vertical vorticity (10^{-4} s^{-1}) at lowest model level for 12 April 1991 simulations. Dashed line is for forecast environmental conditions; thin line for observed temperature and forecast moisture and wind conditions; and thick line for observed temperature and wind conditions with forecast moisture.

the future. Clearly, accurate verification information is needed in order to evaluate the sounding forecast portion of the experiment. The lack of accurate verification information makes determining the reasons for overforecasting or underforecasting by the model difficult at best. Of the overforecast days, the verification sounding from 11 April indicated that the increased moisture put into the sounding turned a good forecast of no severe weather into an overforecast. The only model underforecast sampled with a good verification sounding is 12 April, for which the problem appears to one of model sensitivity. Inhomogeneities in the environment also may have played a role in determining the environment. Since the model begins with a horizontally homogeneous environment, this could be a serious drawback in some situations. Finally, other methods of initializing the model, such as using information from larger-scale models or carrying out multiple forecasts as in the quasi-Monte Carlo approach, need to be explored.

Acknowledgments. This work was done while one of the authors (HEB) held a National Research Council–NOAA Research Associateship. The computer simulations were carried out on the CRAY-2 at the National Center for Supercomputer Applications (NCSA), and software developed at NCSA was used in the analysis of the data. We thank NCSA for allowing us to carry out the simulations at a high priority on their system in order to produce a timely product. The simulations were supported, in part, by NSF Grant ATM-87-00778. Donald Burgess, Charles Crisp, Charles Doswell, and Arthur Witt participated as experimental forecasters from NSSL. David Andra, Fred Beeler, Michael Branick, William Bunting, Michael Foster, Ken Gallant, Dennis Noble, Larry Ruthi, John Victory, and David Wert participated from NWSFO OUN. We thank Dennis McCarthy, NWSFO OUN Area Manager, for his support. We gratefully acknowledge the comments of the three anonymous reviewers for helping to clarify and improve the manuscript.

REFERENCES

Brooks, H. E., 1992: Operational implications of the sensitivity of modelled thunderstorms to thermal perturbations. Preprints, *Fourth Workshop on Operational Meteorology,* Whistler, B.C., Canada, Atmospheric Environment Service/Canadian Meteorological and Oceanographic Society, 398–407. [Reprints available from the author at NOAA/NSSL, 1313 Halley Circle, Norman, OK 73069.]

——, and R. B. Wilhelmson, 1992: Numerical simulation of a low-precipitation supercell thunderstorm. *Meteor. Atmos. Phys.,* **49,** 3–17.

——, 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.

Davies-Jones, R., and H. E. Brooks, 1993: Mesocyclogenesis from a theoretical perspective. *Frontiers in Tornado Research,* C. Church, Ed., *Amer. Geophys. Union,* in press.

——, D. Burgess, and M. Foster, 1990: Test of helicity as a tornado forecast parameter. Preprints, *16th Conf. on Severe Local Storms,* Kananaskis Park, Alberta, Canada, Amer. Meteor Soc., 588–592.

Droegemeier, K. K., 1990: Toward a science of storm-scale prediction. Preprints, *16th Conf. on Severe Local Storms,* Kananaskis Park, Alberta, Canada, Amer. Meteor Soc., 256–262.

Klemp, J. B., and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.,* **35,** 1070–1096.

Murphy, A. H., 1991: Probabilities, odds, and forecasts of rare events. *Wea. Forecasting,* **6,** 302–307.

Rotunno, R., and J. B. Klemp, 1985: On rotation and propagation of simulated supercell thunderstorms. *J. Atmos. Sci.,* **42,** 271–292.

Rust, W. D., R. Davies-Jones, D. W. Burgess, R. A. Maddox, L. C. Showell, T. C. Marshall, and D. K. Lauritsen, 1990: Testing a mobile version of a Cross-Chain Loran Atmospheric (M-CLASS) Sounding System. *Bull. Amer. Meteor. Soc.,* **71,** 173–180.

Straka, J. M., and J. R. Anderson, 1993: Numerical simulations of microburst producing storms: Some results from storms observed during the COHMEX. *J. Atmos. Sci.,* **50,** 1329–1348.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.,* **110,** 504–520.

——, and ——, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shear. *Mon. Wea. Rev.,* **112,** 2479–2498.

Wicker, L. J., and R. B. Wilhelmson, 1990: Numerical simulation of a tornado-like vortex in a high resolution three dimensional cloud model. Preprints, *16th Conf. on Severe Local Storms,* Kananaskis Park, Alberta, Canada, Amer. Meteor Soc., 263–268.