Conditional Tornado Probabilities From Ruc-2 Forecasts

Andrew T. Church

SOARS Protégé
University of New Mexico
Albuquerque, New Mexico

Research Mentor and Co-author:
Dr. Thomas Hamill
Mesoscale and Microscale Meteorology Division
The National Center for Atmospheric Research
Boulder, Colorado

August 11th, 1999

UCAR-SOARS
University Corporation for Atmospheric Research
Significant Opportunities in Atmospheric Research and Science
Sponsored by the National Science Foundation
ABSTRACT

Several previous studies have established statistical relationships between tornadic weather and environmental conditions by associating severe weather reports with rawinsonde observations. Here, we seek (1) to determine whether similar relationships hold when severe weather reports are associated with gridded short-term numerical forecasts, and (2) to develop and demonstrate a probabilistic model to forecast the likelihood a thunderstorm will be tornadic. Severe weather reports and lightning network data from 1 January 1999 through 29 June 1999 were used to classify the weather at a set of Rapid-Update Cycle (RUC-2) grid points into the mutually exclusive and collectively exhaustive events representing increasing severity to the weather. The events were no significant weather, thunderstorms, severe thunderstorms, or tornadoes. RUC-2 forecast convectively available potential energy (CAPE), helicity, and 0-6 km wind shear from these same dates were associated with this gridded classification of the weather. We generally found similar relationships between environmental parameters and storm severity as others have previously documented. Our Bayesian probabilistic model forecasts the likelihood of that a thunderstorm will be tornadic given a certain value of CAPE and helicity (or CAPE and wind shear). This model is shown to do a reasonable job of locating high-threat areas many hours in advance of the severe weather and may be of use as a forecast tool.
1. INTRODUCTION

Forecasts of severe weather have improved substantially over the past few decades due in part to a better understanding of the dynamics of thunderstorms and to an improvement in real-time data sources such as Doppler radars, wind profilers, and the lightning data network. Since 1973, the probability of detecting tornadoes in a region under a tornado watch has increased from 30% to 60% (Grice et al. 1999). Nonetheless, predicting severe thunderstorms other than by extrapolating the motion of existing severe storms is still quite difficult.

Prediction of these phenomena may be improved through the use of more accurate, higher resolution weather forecast models with greater ability to resolve smaller-scale features and orographic effects. Operational and research models continue to be upgraded to higher resolutions and more robust physics. For example, in January, 1996, substantial revisions were made to the Eta's model physics to improve its ability to capture daily and seasonal variations in near-surface air temperature, specific humidity, and wind speed (Rogers et al. 1996; Yucel et al. 1997). In February, 1998, the Eta resolution was increased to 32 km and 45 layers from 48 km and 38 layers. The Rapid Update Cycle system, version 2 (RUC-2) has undergone some significant improvements as well. Horizontal resolution in RUC-2 is 40 km, compared to 60 km for the RUC-1 (Benjamin et al. 1998), and the RUC-2 now uses a more sophisticated treatment of land-surface processes (Smirnova et al. 1997). Many research groups such as the Center for Analysis and Prediction of Storms (CAPS) are also experimenting with real-time short-range forecasts of convection using high resolution numerical models (Xue et al. 1995). The trend toward finer-resolution models will surely continue during the coming decades, and analysis systems and forecast models physics will also likely be improved. Future operational models are likely to be run at resolutions below 10 km and with explicit representation of convection rather than using cumulus parameterizations schemes. This may permit the more accurate treatment of convection in the numerical models (Weisman et al. 1997) and the ability to resolve features such as supercell thunderstorms.
While this new era of modeling brings with it the promise of continued improvements in forecast skill and the ability to soon resolve convective storms in our simulations, chaos theory (Lorenz, 1963; Lorenz 1969) strongly suggests that it will always be difficult or impossible to predict the precise timing and location of these storms beyond a few hours (Islam et al. 1993). Generally, the smaller the scale of the phenomenon, the shorter the range of predictability. Given that precise numerical forecasts of severe weather are likely to remain problematic, a realistic alternative goal is to relate the probability of mesoscale or microscale (Orlanski 1975) severe weather events to larger-scale environmental conditions.

Just what are the important larger-scale environmental conditions? Through observational studies and through theoretical and numerical modeling work, several aspects have been established as being important predictors of severe weather. Miller (1972) and others have suggested the importance of buoyancy parameters; the most typically used now is “CAPE,” or convectively available potential energy, defined by Moncrieff and Miller (1976). Chisholm and Renick (1972) and Fankhauser and Mohr (1977) discussed sounding characteristics typical of single cell, multicell, and supercell thunderstorms, finding among other things that strong vertical wind shear was typical of supercells. Numerical simulations by Klemp and Wilhelmson (1978), Schlesinger (1980), Rotunno and Klemp (1982, 1985), and Weisman and Klemp (1982, 1984, 1986), and Klemp (1987) examined the dependence of storm structure on wind shear and buoyancy through the use of a numerical cloud model. Using judiciously chosen thermodynamic and wind profiles, they demonstrated an ability to simulate storms that were qualitatively similar to those observed and to understand the dynamics of storm-splitting and the deviate motion of supercells. Further, these latter findings helped identify the proper environments for the development of particular types of severe storms (e.g., short-lived storms were associated with low shear and supercell storms associated with high shear).

Storm-relative environmental helicity, or “SREH,” a measure of streamwise vorticity, has also been suggested as an important predictor of supercells and/or tornadic activity (Lilly 1986; Davies-Jones 1984; Davies-Jones et al. 1990). Pictorially, helicity is proportional to the area
swept out on a hodograph relative to the storm-motion vector. Dynamically, helicity in the region of storm inflow induces rotation when ingested into the storm and tilted into the updraft. Davies-Jones et al. (1990), Davies (1993), and Droegemeier et al. (1993) discussed the use of helicity as a forecast parameter for supercell thunderstorms. Recently, Weisman and Rotunno (1999) performed numerical simulations that detail some potential problems interpreting supercell dynamics using helicity. One major problem they discuss is that any quantization of helicity before the occurrence of the actual storm is suspect, since helicity is sensitive to the storm motion vector, which can only be estimated.

Several studies have attempted to establish statistical relationships between environmental conditions and severe weather. Davies and Johns (1993) discussed the relationship of wind shear and helicity to strong and violent tornadoes. Johns et al. (1993) similarly discussed the relationship between the severity of tornadoes and combinations of wind and buoyancy parameters. Rasmussen and Blanchard (1998), hereafter "RB98," developed statistical relationships between thunderstorm severity and environmental parameters as measured by nearby rawinsondes. When convection occurred (> 10 cloud-to-ground lightning strikes in the area), the convection was classified as non-supercellular, supercellular without tornadoes, or tornadic. Environmental parameters were associated with the sample thunderstorm classification using data from a relatively nearby 0000 UTC sounding. Parameters such as mean shear, 0-6 km shear, helicity, CAPE, and combinations thereof were examined. Individual parameters were generally shown to discriminate less effectively between the three thunderstorm classifications than combinations of parameters such as CAPE and 0-6 km shear or CAPE and helicity.

A recent study conducted by Stensrud et al. (1997) suggested mesoscale model output may be used to discriminate between tornadic and nontornadic thunderstorms. The simulated events included tornadic supercell thunderstorm outbreaks associated with numerous reports of tornadoes, bow-echo storms that had several tornadoes, and isolated supercell events. The study found that helicity greater than 100 $m^2s^{-2}$ is useful as a threshold for determining regions where
supercell thunderstorms were likely to form. High bulk Richardson number wind shears suggested low-level mesocyclogenesis was likely.

If tornado likelihood is to be forecast many hours prior to tornado occurrence, environmental conditions from numerical model forecasts such as in Stensrud et al. (1997) will be required. To this end, we have collected half a year's worth of RUC-2 analyses and forecasts and severe weather observations. Our aims are (1) to determine whether some relationships between severe weather and environmental parameters are similar when diagnosed from model analyses and 12-h forecasts; and (2) to develop a probabilistic forecast model for the conditional probability of tornadoes given that a thunderstorm occurs based on RUC-2 12-h forecasts.

Based on the collective wisdom of many previously cited studies, we will focus primarily on combinations of CAPE and 0-6 km shear or CAPE and helicity. These parameters (and many others) are considered when the National Weather Service's Storm Prediction Center (SPC) generates their severe weather outlooks (Johns and Doswell 1992; Doswell et al. 1993). Our model may be considered a simple, automated prediction algorithm that uses the most widely accepted precursors to severe weather. We acknowledge that there are many other potentially crucial factors relating to severe storm development, including the strength of capping inversions and specific triggering mechanisms such as outflow boundaries or drylines (e.g., Markowski et al. 1998; Rasmussen et al. 1999). This prototype study should be regarded as a proof-of-concept; we will not consider such additional effects here but acknowledge the wisdom of considering them in later studies, as well as the wisdom of collecting and analyzing many years' worth of data.

The paper will be organized as follows. Section 2 describes the data and methodology used in this 6-month climatological analysis. The method for categorizing storms as tornadic, severe, or ordinary will be described as well as our methods for calculating shear and helicity parameters. Section 3 compares the relationships between storm types and environmental conditions to those found in RB98. Section 4 then describes the probabilistic forecast model we developed. Section 5 presents some prototype tornado probability forecasts using this model. Section 6 concludes.
2. DATA AND METHODOLOGY

2.1 Classification of observed weather at model grid points

In order to match severe weather with gridded forecast data, we developed an algorithm to classify the weather at RUC-2 model grid points. Because of the paucity of reported data in the west, the climatology was based only on a "masked" subset of points in the central and eastern U.S. (shaded points in Fig 1). The severe reports analyzed here originated from the SPC's "rough log" from January 1, 1999 through June 30, 1999. The log is a preliminary database based on local storm reports SPC provides with few quality control checks in place and no indication of the severity of the tornado. No attempt was made to verify the reports. Only observational data from 2200 to 0200 UTC were used; it was assumed that this time window could reasonably be associated with numerically forecast or analyzed conditions at 0000 UTC. Slightly more than half of the day's tornadoes occurred within this 4-hour window. The National Lightning Detection Network, operated by Geomet Data Services, Inc., provided the cloud-to-ground (CG) lightning data.

The composited severe report and lightning databases were used to classify the weather at each RUC-2 grid point within the masked area into one of four mutually exclusive and collectively exhaustive categories representing the maximum severity of the weather in that grid box. The categories were (1) no significant weather, (2) ordinary thunderstorm, (3) severe weather (high winds or large hail), or (4) tornado. No significant weather event took place in the grid box surrounding the grid point if no severe reports, no tornado reports, and less than 10 CG lightning strikes occurred (the choice of 10 is admittedly somewhat arbitrary). We classify a grid point as a thunderstorm if more than 10 CG lightning strikes occurred in the grid box between 2200 UTC and 0200 UTC but no severe weather nor tornadoes were reported. The grid point was classified as severe if a severe report occurred within the grid box between 2200 UTC and 0200 UTC. The grid point was classified as tornadic if there was an observation of a tornado in that grid box during that time. An example of the classification process is provided in Figs. 2
and 3. Figure 2 shows the SPC severe weather and tornado reports that occurred within the 4-hour time window of 0000 UTC 22 January 1999. Figure 3 shows the gridded classification of the weather at the RUC-2 grid points using the SPC reports combined with lightning data.

Our classification is less than ideal. Since concepts like helicity were conceived with supercellular convection in mind, ideally we would like to carefully discriminate between supercellular convection and other convection. However, we did not have the time to carefully examine radar data to discriminate supercellular convection from other convection, nor did we have quality-controlled severe weather reports coincident with the RUC-2 archived data. Hence, we could not refine our categorization of severe weather much more carefully than we already have done without drastically increasing the complexity of data processing. The influence of potentially incorrect severe weather reports, the use of only half a year’s data, and the mixing of supercellular with non-supercellular tornadoes should be considered when interpreting our results.

2.2 *Defining CAPE, Shear, and Helicity from RUC-2 data*

We now discuss how environmental parameters were determined from the RUC-2. Less precise relationships between severe weather and environmental parameters should be expected when using RUC-2 model data (as compared to appropriate pre-convective proximity soundings). Model-derived parameters inevitably contain some error, potentially obscuring true relationships. However, given that forecasters may want to use 12-h RUC-2 forecasts for diagnosing the threat of severe weather, we seek to determine whether previously documented relationships hold when using gridded analyses or model forecasts.

To this end, RUC-2 0000 UTC analyses and 12-h forecast from 1200 UTC condition were used to define CAPE, wind shear, and helicity at model grid points inside the mask. Helicity was calculated using 0-3 km winds and following the Galilean-invariant methodology for calculating storm-relative motion described in Bunkers et al. (1999). “Wind shear” hereafter will refer to the wind speed difference between 0 and 6 km above ground level (AGL), or what RB98
call "BL-6 km shear." We examined shear calculated in other manners; results were qualitatively similar.

If convection is occurring at a forecast grid point at 0000 UTC, its forecast sounding may be unrepresentative because there may already have been convection and release of CAPE. To address this, we have developed our own proximity sounding method, as illustrated in Fig. 4. If there is rain occurring at a model grid point, a set of adjacent grid points are examined to determine if any of them are rain free. This set of adjacent grid points is determined by first finding the grid point that is 160 km upstream of the rainy grid point, with the upstream direction defined using surface winds. A $9 \times 9$ array of grid points centered on this upstream grid point is considered, and the nearest rain-free grid point to the original rainy grid point is located. If the CAPE at this rain-free grid point is greater than the CAPE at the rainy grid point, then the CAPE, shear, and helicity values from the rain-free grid point replace the values at the rainy point. If the rainy point's CAPE value is higher, it retains its CAPE, shear, and helicity. This algorithm is, of course, less than ideal. If we had RUC-2 data available at very high temporal resolution, then ideally we would simply use the RUC-2 conditions at the grid point immediately preceding the onset of convection. However, since we lacked this high temporal-resolution data, this approximate algorithm was chosen to address some of the issues of unrepresentative soundings.

3. RELATIONSHIPS BETWEEN WEATHER AND ENVIRONMENTAL PARAMETERS

Here, we analyze the relationship between the classified weather and the environmental parameters using statistics similar to those used in RB98. This will permit ready comparison to see if the relationships are robust despite our use of analyzed and forecast data (as opposed to their use of raobs), and given some improved or slightly different analysis methods. We will examine the relationships between the weather classification and CAPE, helicity, 0-6 km shear, the
energy-helicity index (Hart and Korotky 1991; Davies 1993), and the vorticity-generation parameter, (Rasmussen and Wilhelmson 1983).

Consider first the scatterplots of weather as a function of RUC-2 12-h forecast CAPE and shear (Figs. 5 a-c) and CAPE and helicity (Figs. 6 a-c). As shown in the scatterplots, there are drastically many more points with ordinary thunderstorms than with tornadic thunderstorms, emphasizing the rarity of tornadic events. However, the frequency of tornadoes relative to ordinary thunderstorms is much more pronounced at higher CAPE and shear values than at lower CAPE and shear values. That is, at lower CAPE and shear, ordinary thunderstorm counts overwhelm severe counts and tornado counts, but at higher CAPE and shear, the relative proportions of tornadic, severe, and ordinary thunderstorms are much more similar.

The relationships can be better quantified with box and whiskers plots, as in RB98. Plots for the severe weather classification as a function of CAPE alone are shown in Figs. 7 a-b. As shown, when considering CAPE alone, CAPE appears to rather effectively discriminate between the different types of weather; ordinary thunderstorms are associated with much lower CAPE than are severe thunderstorms, which in turn are associated with lower CAPE than tornadic thunderstorms. Analysis data appears to discriminate more effectively than forecast data, as expected, given its lower error. Note that our classification here seems to show a stronger relationship between the severity of the weather and CAPE than in RB98 and other studies. The extent to which this is due to the different proximity sounding methods used or the different data (RUC-2 gridded data vs. raobs, and our January-June data rather than their full year's data) is not clear.

Box and whiskers plots for 0-6 km shear are shown in Figs. 8 a-b. Shear alone does not appear to be a particularly effective discriminator, though there does appear to be a shear threshold for tornadoes; 90 percent of the tornadoes occurred when shear was $\sim 15 \, ms^{-1}$ shear or greater. No such threshold at a relatively high shear existed when considering ordinary thunderstorms.

Helicity (Figs. 9 a-b) does not appear to be as effective a discriminator here as was indicated in RB98. Nonetheless, as with shear, there appears to be a lower limit to helicity below which tornadoes are extremely unlikely. Note, though, that our classification methodology dif-
fers from that used in RB98. For this paper, helicity was calculated with a different, presumably improved method for calculating storm-relative motion (Bunkers et al. 1999). Further, RB98 used radar data and quality-controlled severe storm reports to classify events as ordinary thunderstorms, supercells without significant tornadoes, or significant tornadoes (F-2 or greater). Our classification of a grid point as tornadic included weaker tornadoes and tornadoes associated with non-supercellular thunderstorms, so helicity may not be as relevant as a discriminator.

As in RB98, we examined combinations of parameters, namely, the energy-helicity index (EHI; Hart and Korotky 1991; Davies 1993) and the vorticity-generation parameter (VGP; Rasmussen and Wilhelmson 1998). EHI is defined as

$$EHI = \frac{(CAPE)(SREH)}{1.6 \times 10^5}.\quad (1)$$

Previously, EHI > 2.0 was suggested as indicating a large probability of supercells. The VGP is defined as

$$VGP = (S)\sqrt{CAPE}, \quad (2)$$

where $S$ is the mean shear (the hodograph length divided by the depth). As explained in RB98 and references therein, the VGP approximates the rate of conversion of horizontal to vertical vorticity.

Figures 10 a-b and 11 a-b present box and whiskers plots for EHI and VGP, respectively. As with RB98, the EHI appears to discriminate relatively well between the three storm severities, while the VGP is slightly less effective as a discriminator.

Overall, the relationships demonstrated in this section suggest that RUC-2 forecast data does have some ability to discriminate between ordinary, severe, and tornadic storms despite the fact that its soundings will have some errors.

4. MODEL FOR TORNADO PROBABILITIES

4.1 Model Design

We now turn our attention to how the forecast CAPE and shear information may be used to model the likelihood of tornado occurrence. A similar model was developed for CAPE and
helicity and for RUC-2 analyses using the same principles discussed below, but for brevity, we describe the process only for CAPE and 0-6 km shear from RUC-2 forecasts. As shown previously, these forecast parameters show some ability to discriminate amongst types of thunderstorms, and combinations of parameters were even more effective. Hence, we seek to build a probabilistic model to determine the conditional probability that a thunderstorm will be tornadic, given that a thunderstorm occurs in the RUC-2 grid box. Such a model could be used as a forecast tool to quickly determine areas of enhanced risk. Our models admittedly neglect other potentially important parameters such as convective inhibition or the effects of low-level boundaries (Markowski et al. 1998; Rasmussen et al. 1999).

We could make a crude estimate of conditional probability of tornadoes at a given CAPE and shear; we could count the number of dots around a particular CAPE and shear value on Fig. 5c and divide by the total number of dots around that CAPE and shear in Figs. 5a, b, and c. We seek, however, to be a bit more careful in our estimation of probabilities.

The starting point for an improved model is Bayes’ Rule (Wilks 1995). In this model, let $B$ be the compound event that 0-6 km shear $Sh$ and CAPE have certain forecast values. For example, $B$ might be the event that $Sh \geq 20 \text{ m s}^{-1}$ and $\text{CAPE} \geq 3000 \text{ J kg}^{-1}$. Let $T$ be the event that a RUC-2 grid box is classified as tornadic; let $S$ be the event that it is classified as severe; and let $R$ be the event that it is classified as an ordinary thunderstorm. Then, Baye’s rule states

$$P(T|B) = \frac{f(B|T)P(T)}{f(B)},$$

(3)

where

$$f(B) = \sum f(B|T)P(T) + f(B|S)P(T) + f(B|R)P(R).$$

(4)

Here, $P(T) = n_T/(n_T + n_S + n_R)$, $P(S) = n_S/(n_T + n_S + n_R)$, and $P(R) = n_R/(n_T + n_S + n_R)$. $n_T$, $n_S$, and $n_R$ indicate the total number of grid boxes classified as either tornadoes, severe thunderstorms, or ordinary thunderstorms. That is, $n_T$, $n_S$, and $n_R$ are the total number of dots in Figs. 5a-c, respectively. The probability density functions $f(B|T)$, $f(B|S)$, and $f(B|R)$ represent the probability density of a particular CAPE and shear value given that a tornado, severe thunder-
storm, or ordinary thunderstorm occurs, respectively. Qualitatively, they are related to the local density of dots on each of the diagrams relative to the density in other areas of the diagram.

We thus need to estimate a probability density function for tornadoes as a function of CAPE and shear (and estimates for severe and ordinary thunderstorms as well). We will constrain the possible values of CAPE and shear to realistic maximum and minimum values and estimate densities within these parameter ranges. More specifically, we seek to estimate \( f(B|T) = f((\text{CAPE}, Sh)|T) \), where \( 0 \leq \text{CAPE} < 6000 \, J \, kg^{-1} \) and \( 0 \leq Sh < 100 \, m \, s^{-1} \). Probabilities integrate to 1.0 over this domain. \( \int \int f((\text{CAPE}, Sh)|T) \, d\text{CAPE} \, dSh = 1 \). \( f(B|S) \) and \( f(B|R) \) were similarly determined.

Estimating \( f(B|T) \), \( f(B|S) \), and \( f(B|R) \) with parameteric probability density functions such as the gamma distribution or normal distributions fitted to power-transformed data (Wilks 1995) produced unsatisfying results (not shown). Hence, we resorted to estimating \( f(B|T) \), \( f(B|S) \), and \( f(B|R) \) using non-parametric density estimation techniques (Silverman, 1986). The details of this technique are supplied in the appendix.

Contours of our selected probability density estimates are overlaid on the scatterplots in Figs. 5-6 a-c. These probability density estimates are then used in conjunction with RUC-2 forecast parameters using (3) and (4) to determine the conditional probability a thunderstorm will be tornadic. An illustration of this model's estimate of conditional tornado probability as a general function of CAPE and shear and CAPE and helicity are shown in Figs. 12 a-b. As expected, the probability of a tornado occurring generally increases with increasing CAPE and shear. Even with the relatively strong smoothing used to generate the density estimates, the probabilities have multiple maxima that are almost certainly unrealistic. We would expect that if we collected a longer sample of data, these curves would become smoother and not exhibit such multiple maxima.

5. RESULTS FROM PROBABILISTIC MODEL
We now provide some simple illustrations of conditional tornado probability forecasts generated from this model. We shall examine tornado forecasts for 0000 UTC on three days: 22 Jan 1999, 16 April 1999, and 4 May 1999. The first and third of these dates are days with significant tornado outbreaks in Arkansas (Fig. 3) and Oklahoma (Fig. 13). The remaining case day was randomly selected from the many days where there was no severe weather. Results are shown for this case day to demonstrate the ability of the model to avoid false alarms.

It is important to develop the probabilistic model with a different data set than is used to evaluate it. To this end, when producing forecasts for 22 January 1999, for example, this day’s data points were excluded from the data set used to develop probability estimates, with similar exclusion for the other case days.

Tornado probability forecasts valid at 0000 UTC 22 Jan 1999 are shown in Figs. 14 a-b, with the first panel showing the forecast based on CAPE and shear and the second panel based on CAPE and helicity. Dotted areas indicate regions where probabilities may not be reliable due to CAPE and shear or helicity values being outside the range in the training data, typically due to unusually large values of one of the predictors. The model has a maxima of probability in eastern Texas with an axis of elevated probability through north central Arkansas. Why is the location of the maximum of tornado probability misforecast? In this case, we do not believe it is actually so much a problem with the model as the problem with the RUC-2 data feeding the model. Consider Figs. 15 a-b, which show the (slightly smoothed) RUC-2 12-h forecast and analyzed CAPE at 0000 UTC 22 Jan 1999. As shown, the RUC-2 misforecasts the location of maximum CAPE, and the actual location of tornadoes is actually nearly coincident with the region with largest CAPE. This illustrates that the model cannot be expected to be any better than the forecast data fed into it; specifically, this model cannot correct for errors in position or intensity of the relevant CAPE, shear, and helicity features. As has been demonstrated for with operational precipitation forecasts, perhaps these probability forecasts can be judiciously adjusted by humans with a knowledge of the forecast model’s typical random and systematic errors.
Probability forecasts for the quiescent day, 0000 UTC 16 April 1999, are shown in Figs. 16 a-b. Note there are no regions with highly elevated probabilities of tornadoes. The dotted regions in Fig. 16a are a region in the core of the jet with no CAPE but very large shear values.

The forecasts for the Oklahoma City tornado outbreak at 0000 UTC 4 May 1999 are shown in Figs. 17 a-b. The maximum for tornado probability was forecast for the panhandle of Texas, but there was also a relative maxima through Oklahoma and Kansas, coincident with the area struck by tornadoes. There was some convection associated with very large CAPE values in south-central Texas that day. However, as indicated by the dots on the figure, the high tornado probabilities are probably untrustworthy due to an insufficient number of other cases with similar parameter values. Further, 0000 UTC soundings (not shown) indicated that much of that area was strongly capped. This case result suggests that this forecast method might benefit both from re-developing the model using a longer training data set and refining it by also considering convective inhibition as an additional predictor.

6. CONCLUSIONS

This paper described results from associating a classification of thunderstorm severity (ordinary, severe, and tornadic) based on severe weather reports and lightning observations with RUC-2 analyses and forecasts. The purpose was both to compare and contrast these model-based results with the raob-based results such as in RB98 and to develop a probabilistic forecast model for conditional tornado likelihood. This model specifies the conditional probability that a tornado will occur given that a thunderstorm occurs and given that certain RUC-2 CAPE and shear (or CAPE and helicity) values are forecast.

Our comparison with RB98 results showed that we achieved better discrimination of storm types using CAPE than they did, but worse discrimination using helicity. Differences may have been due to either slightly different classification methods (they used quality-controlled severe reports and classified a sounding as tornadic only if there was an F-2 or greater tornado) or due
to our use of gridded analyses and forecasts rather than raobs, or due to different training periods (ours included little summertime data).

Our probabilistic model was developed in a Bayesian framework. A few select cases demonstrated that the model does a reasonable job of predicting areas of enhanced risk for tornadoes. These results also suggest that when the probabilistic model misforecasts tornado likelihood, it is sometimes due to the RUC-2 forecast parameters being in error rather than a problem with the model design. A probabilistic model like this one may be of use as a supplemental guidance tool for severe storms forecasters.

Ideally, it would be preferable to include other forecast parameters such as convective inhibition into this probabilistic model, to use quality-controlled severe storm reports, and to develop a model using a much longer training period than the six months used here.

7. APPENDIX: DETAILS OF DENSITY ESTIMATION.

Nonparametric density estimation techniques allow a user to fit a complicated probability distribution to scattered samples of data. The technique is called “nonparametric” because no prior parameterized form of a probability distribution (e.g. Gaussian, or gamma, or Weibull) is assumed. An extensive description of these techniques is beyond the scope of this paper; the reader is referred to Silverman (1986) for a readable introduction to the topic.

To start, let \((U', V')\) be a continuous random vector with components representing values of CAPE and shear or CAPE and helicity. We seek to estimate the probability density function (pdf) \(f_{U',V'}(u',v')\) at a regular set of points in a bounded domain for a given weather classification, with bounds \(0 \leq \text{CAPE} \leq 6000 \ J kg^{-1}, 0 \leq \text{Sh} \leq 100 \ ms^{-1}\), and \(-100 < \text{helicity} \leq 1200 \ m^2 s^{-2}\). Note that this is the same pdf as \(f(B|\cdot)\) in the notation of equation (4) in the main text. For example, let us assume we are trying to estimate \(f(B|R)\) in equation (4).

For computational simplicity we normalize the sample data and coordinate system so most samples fall within [0,1], i.e., \(U = (U' - U'_{\text{min}})/(U'_{\text{max}} - U'_{\text{min}})\) and \(V = (V' - V'_{\text{min}})/(V'_{\text{max}} - V'_{\text{min}})\). Here, \(U'_{\text{min}} = 0 \ J kg^{-1}, U'_{\text{max}} = 6000 \ J kg^{-1}, V'_{\text{min}} = 0 \ ms^{-1}\) and \(V'_{\text{max}} = 100 \ ms^{-1}\)
for wind shear, and $V'_m = -100 \ m^2 s^{-2}$ and $V'_{m_{max}} = 1200 \ m^2 s^{-2}$ for helicity. A necessary step is to thus determine the estimated pdf $\hat{f}_{U,V}(u, v)$ in the normalized coordinate system.

Since the areal extent of the domain for the pdf is shrunk (for CAPE and shear) by $6000 \times 100 = 600,000$, and since probability density is assumed to integrate to 1.0 over both domains, i.e., $\iint \hat{f}_{U,V}(u', v') \ du' \ dv' = \iint \hat{f}_{U,V}(u, v) \ du \ dv = 1$,

$$\hat{f}_{U',V'}(u', v') = \frac{\hat{f}_{U,V}(u, v)}{600,000}. \quad (a1)$$

Nonparametric density estimates in this normalized coordinate system were found to be highly inaccurate due to the data being strongly skewed and bounded on the left, as shown in the scatterplot data in Figs. 5-6. Hence, we seek a transformation to yet another coordinate system to make the data more normally distributed (Wilks 1995). Power transformations are of the form

$$x = \frac{u^{\lambda_c} - 1}{\lambda_c} = h_1(u) \quad (a2)$$

for CAPE, and

$$y = \frac{v^{\lambda_{sh}} - 1}{\lambda_{sh}} = h_2(v) \quad (a3)$$

for shear and helicity. Reasonable choices for the exponents were found to be $\lambda_c = 0.5$, and $\lambda_{sh} = 0.7$ or $\lambda_{sh} = -0.1$ for wind shear and helicity, respectively.

The joint pdf in the original, normalized coordinate system is related to the pdf in the power-transformed coordinate system by

$$\hat{f}_{U,V}(u, v) = \hat{f}_{X,Y}(x, y) |J|, \quad (a4)$$

where $| \cdot |$ represents the absolute value (Casella and Berger 1990). The Jacobian $J$ is

$$J = \det \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u} \frac{\partial y}{\partial v}. \quad (a5)$$

The derivatives of the power transformations are necessary for the transformation of the pdf between coordinate systems. From equations (a2) and (a3), $\frac{\partial x}{\partial u} = u^{\lambda_c - 1}$; $\frac{\partial x}{\partial v} = 0$; $\frac{\partial y}{\partial u} = 0$; and $\frac{\partial y}{\partial v} = v^{\lambda_{sh} - 1}$. 

17
Our approach to estimating $\hat{f}_{X,Y}(x, y)$ is to first estimate the relative shape of the pdf as it varies across the power-transformed computational domain and then later to normalize to ensure probability density integrates to 1.0. That is, step 1: determine an unnormalized $\hat{f}_{X,Y}^*(x, y)$, as described below, and step 2:

$$\hat{f}_{X,Y}(x, y) = \frac{\hat{f}_{X,Y}^*(x, y)}{\int_{h_2(V_{\text{max})}}^{h_2(V_{\text{max}})} \int_{h_1(U_{\text{max}})}^{h_1(U_{\text{max}})} \hat{f}_{X,Y}^*(x, y) \, dx \, dy}$$

(a5)

We determine $\hat{f}_{X,Y}^*(x, y)$ using a variant of a kernel density estimation technique with a Cressman (1950) kernel. An approach equivalent to the method of Fukunaga (Silverman 1986) was used to stretch the kernel in the direction of correlations of the data. We found that $\sigma(x)/\sigma(y) \approx 2$, and $\text{Cov}(x,y) \approx 0.0$, indicating that the kernels ought to be stretched out twice as much in the x-direction as in the y-direction. Accordingly, we assume

$$\hat{f}_{X,Y}^*(x, y) = \frac{1}{nh^2} \sum_{i=1}^{n} K(d),$$

(a7)

where

$$d = \sqrt{(x - X_i)^2 + 2(y - Y_i)^2}.$$  

(a8)

and

$$K(d) = \begin{cases} \frac{h^2 - d^2}{h^2 + d^2}, & \text{if } d < h; \\ 0, & \text{otherwise}. \end{cases}$$

Here, $(X_i, Y_i)$ are a sample vector with components (CAPE, shear) or (CAPE, helicity), and there are n samples for this weather classification. h is a window width controlling the amount of smoothing. How do we choose h? An assumption underlying density estimation is that each sample point is independent. This assumption is not valid here; e.g., many points in Figs. 5 a-c may be clustered together, representing a single outbreak on a given day. Hence, determining the optimal smoothing through conventional methods such as cross-validation (Silverman 1986) produced density estimates that were not smooth enough. In the end, we decided to select window widths manually, based on a smoothness we thought would be relatively realistic if we had a much larger sample size.
Once we have $\hat{f}_{X,Y}(x, y)$ from (a7), we get $\hat{f}_{X,Y}(x, y)$ from (a6) and then use equation (a4) to get $\hat{f}_{U,V}(u, v)$. Finally, we apply (a1) to produce the desired $\hat{f}_{U',V'}(u', v')$, i.e., the pdfs overlaid on the scatterplots in Figs. 5 and 6.

8. ACKNOWLEDGMENTS

The authors would like to thank Greg Thompson (NCAR/RAP) and Tracy Smith (NOAA/FSL) for their help in obtaining RUC-2 data. Tressa Kane and Randy Bullock (NCAR/RAP) provided assistance in obtaining and decoding the lightning network data. Mike Kay at the SPC kindly supplied the rough log data. Doug Nychka (NCAR/GSP) is thanked for his expertise in the development of nonparametric density estimates. Morris Weisman (NCAR/MMM) and Jeff Trapp (NOAA/NSSL) are thanked for substantive advice and for their informal reviews of this manuscript.
REFERENCES


Rogers, E., and others, 1996: Changes to the operational early eta analysis/forecast system at the National Centers for Environmental Prediction. *Wea. and Forecasting*, 11, 391-413.


FIGURE CAPTIONS

Figure 1. Mask showing area where RUC-2 forecasts and observations are compared in this study (shaded area).

Figure 2. Severe weather reports from the SPC rough log for the 4-hour period centered on 0000 UTC 22 Jan 1999.

Figure 3. Gridded analysis of weather type at RUC-2 grid points for the same 4-hour period as in Fig. 3.

Figure 4. Illustration of the method used for selecting an appropriate proxy grid point for a proximity sounding if it is raining at the original grid point. Shaded area represent grid points where rain has accumulated during the past hour. Arrow indicates the direction of the surface wind. Boxed point denotes the center of search region. Solid line encompasses all grid points in search region. Filled circle indicates grid point selected as the nearest rain-free grid point upstream of the original grid point.

Figure 5. Scatterplots of weather types as a function of CAPE and 0-6 km 12-h forecast wind shear from the RUC-2 model. Contours of fitted probability density functions are overplotted (10th, 30th, 50th, 70th, and 90th percentiles). (a) Ordinary thunderstorm, (b) severe thunderstorm, and (c) tornadic thunderstorm.

Figure 6. As in Fig. 5, but for CAPE versus helicity.

Figure 7. Box and whiskers plots for CAPE for the three storm types. Box top, middle, and bottom indicate the 75th, 50th, and 25th percentiles of the empirical distributions, and the top and bottom ends of the whiskers denote the 90th and 10th percentiles, respectively. (a) RUC-2 analysis data, and (b) 12-h RUC-2 forecast data.

Figure 8. As in Fig. 7, but for 0-6 km wind shear.

Figure 9. As in Fig. 7, but for helicity.

Figure 10. As in Fig. 7, but for the energy-helicity index.
Figure 11. As in Fig 7, but for the vorticity-generation parameter.

Figure 12. Conditional probabilities of the occurrence of a tornado given that a thunderstorm occurs as a function of RUC-2 12-h forecast using parameters (a) CAPE and 0-6 km shear, and (b) CAPE and helicity. Areas without contours were not sufficiently populated to attempt a probability estimate.

Figure 13. Severe weather reports from the SPC rough log for the 4-hour period centered on 0000 UTC 4 May 1999.

Figure 14. Forecast conditional probabilities of the occurrence of a tornado given that a thunderstorm happens, for the forecast valid at 0000 UTC 22 Jan 1999. Compare to observations in Fig. 2. (a) Probabilities using CAPE and shear, and (b) using CAPE and helicity. Dotted areas indicate regions where probabilities should be regarded with suspicion because parameter values were outside the range of values used to develop the equations.

Figure 15. RUC-2 12-h forecast (a) and analyzed (b) CAPE valid at 0000 UTC 22 Jan 1999.

Figure 16. As in Fig. 14, but for the date 0000 UTC 16 Apr 1999.

Figure 17. As in Fig. 14, but for the date 0000 UTC 4 May 1999. Compare to observations in Fig. 13.
Figure 1. Mask showing area where RUC-2 forecasts and observations are compared in this study (shaded area).
Figure 2. Severe weather reports from the SPC rough log for the 4-hour period centered on 0000 UTC 22 Jan 1999.
Figure 3. Gridded analysis of weather type at RUC-2 grid points for the same 4-hour period as in Fig. 3.
Figure 4. Illustration of the method used for selecting an appropriate proxy grid point for a proximity sounding if it is raining at the original grid point. Shaded area represent grid points where rain has accumulated during the past hour. Arrow indicates the direction of the surface wind. Boxed point denotes the center of search region. Solid line encompasses all grid points in search region. Filled circle indicates grid point selected as the nearest rain-free grid point upstream of the original grid point.
Figure 5. Scatterplots of weather types as a function of CAPE and 0-6 km 12-h forecast wind shear from the RUC-2 model. Contours of fitted probability density functions are overplotted (10th, 30th, 50th, 70th, and 90th percentiles). (a) Ordinary thunderstorm, (b) severe thunderstorm, and (c) tornadic thunderstorm.
Figure 6. As in Fig. 5, but for CAPE versus helicity.
Figure 7. Box and whiskers plots for CAPE for the three storm types. Box top, middle, and bottom indicate the 75th, 50th, and 25th percentiles of the empirical distributions, and the top and bottom ends of the whiskers denote the 90th and 10th percentiles, respectively. (a) RUC-2 analysis data, and (b) 12-h RUC-2 forecast data.
Figure 8. As in Fig. 7, but for 0-6 km wind shear.
Figure 9. As in Fig. 7, but for helicity.
Figure 10. As in Fig. 7, but for the energy-helicity index.
Figure 11. As in Fig 7, but for the vorticity-generation parameter.
Figure 12. Conditional probabilities of the occurrence of a tornado given that a thunderstorm occurs as a function of RUC-2 12-h forecast using parameters (a) CAPE and 0-6 km shear, and (b) CAPE and helicity. Areas without contours were not sufficiently populated to attempt a probability estimate.
Figure 13. Severe weather reports from the SPC rough log for the 4-hour period centered on 0000 UTC 4 May 1999.
Figure 14. Forecast conditional probabilities of the occurrence of a tornado given that a thunderstorm happens, for the forecast valid at 0000 UTC 22 Jan 1999. Compare to observations in Fig. 2. (a) Probabilities using CAPE and shear, and (b) using CAPE and helicity. Dotted areas indicate regions where probabilities should be regarded with suspicion because parameter values were outside the range of values used to develop the equations.
12 hr forecast from 1200 UTC on 21 Jan 1999, CAPE (J/kg)

Figure 15. RUC-2 12-h forecast (a) and analyzed (b) CAPE valid at 0000 UTC 22 Jan 1999.
Figure 16. As in Fig. 14, but for the date 0000 UTC 16 Apr 1999.
Figure 17. As in Fig. 14, but for the date 0000 UTC 4 May 1999. Compare to observations in Fig. 13.