

Evaluation of Alberta Hail Growth Model Using Severe Hail Proximity Soundings from the United States

RYAN JEWELL

Storm Prediction Center, Norman, Oklahoma

JULIAN BRIMELOW

Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada

(Manuscript received 13 November 2008, in final form 5 June 2009)

ABSTRACT

A one-dimensional, coupled hail and cloud model (HAILCAST) is tested to assess its ability to predict hail size. The model employs an ensemble approach when forecasting maximum hail size, uses a sounding as input, and can be run in seconds on an operational workstation. The model was originally developed in South Africa and then improved upon in Canada, using high quality hail verification data for calibration. In this study, the model was run on a spatially and seasonally diverse set of 914 modified severe hail proximity soundings collected within the contiguous United States between 1989 and 2004. Model output was then compared to the maximum observed hail size for each proximity sounding. Basic verification statistics are presented, showing that the HAILCAST model exhibits considerable skill that can be of use to the operational severe weather forecaster.

1. Introduction

Hailstorms are potentially dangerous and can cause property damage as well as injuries. A prime example of this is the “Mayfest” hailstorm that occurred over north Texas on 5 May 1995. During this event, over 100 people were injured by hail, most of whom were caught out in the open during an outdoor festival (Edwards and Thompson 1998). Even within the relative safety of a vehicle or home, injuries can still result as large stones can penetrate through windows and roofs (Morris and Janish 1996). In addition to bodily injury, hailstorms account for about \$1.3 billion in crop damage and around \$1.3 billion in property damage annually (NSSL 2008; Changnon 1972). Hail is one of the most common types of severe convective weather in the United States, yet hail forecasting has received relatively little attention as compared to that given to tornadoes and damaging winds. While the overall threat to life from severe hail is miniscule in comparison, deaths

can occur (NCDC 2000), and the threat to property can be great.

Forecasters have few reliable tools to aid them in predicting the maximum expected hail size. Radar-based methods of hail detection exist (Donavon and Jungbluth 2007) but are only useful once storms have formed. Historically, attempts at forecasting hail size have focused on various measures of convective available potential energy (CAPE) and temperature levels aloft based on large-scale environment data. One of the first methods used to forecast hail size was developed by Fawbush and Miller (1953), and was based empirically on the buoyancy measured between the convective condensation level (CCL) and the -5°C level. Later, a change was made to incorporate the height of the freezing level (FZL; Miller 1972). Foster and Bates (1956) developed a similar method for forecasting hail size, based on the vertical velocity from the buoyancy between the level of free convection (LFC) and the parcel at the -10°C level, and the predicted terminal velocities of hailstones. A few years later, Renick and Maxwell (1977) developed a nomogram relating maximum hail size to maximum updraft velocity and the temperature at that level during the Alberta Hail Project. More recently, Moore and Pino (1990)

Corresponding author address: Ryan Jewell, Storm Prediction Center, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: ryan.jewell@noaa.gov

constructed a method using the buoyancy derived from integration between the CCL and the -10°C level within the theoretical updraft, in combination with more robust melting calculations.

These methods frequently forecasted unrealistically large hail sizes or showed little ability to delineate between small and large hail (e.g., Doswell et al. 1982). Thermodynamic parameters such as the height of the FZL, wet-bulb zero (WBZ), and CAPE remain in use today, despite their limited utility in forecasting maximum hail size (Edwards and Thompson 1998; Kitzmiller and Breidenbach 1993). A major reason for this is that hail growth is very complex and dependent on storm-scale processes and parameters that are not easily observed. Consequently, any ingredients-based mesoscale forecasting methods that are used to forecast severe weather prior to convective development will result in poor hail size forecasts.

A one-dimensional coupled hail and cloud model, called HAILCAST, has been developed to predict the maximum expected hail diameter D at the surface (Brimelow et al. 2002). HAILCAST has been implemented and tested at the Storm Prediction Center (SPC) over the past 6 yr with promising results.

The SPC forecasts the probability of severe hail ($D \geq 0.75$ in. or 1.91 cm) in the day 1 convective outlook product, as well as the probability that hail will exceed a threshold of 2.0 in. [5.1 cm, which is defined as significant or SIG hail; see Hales (1988)]. In addition, maximum hail diameter is explicitly forecast in tornado and severe thunderstorm watches. Therefore, we tested the model's ability to forecast hail size, and its ability to successfully delineate between nonsignificant severe hail (NON-SIG, $0.75 \leq D < 2.0$ in.) and SIG hail environments. This work is the first time HAILCAST has been tested extensively in warm, moist environments characterized by high CAPE and strong vertical wind shear.

At the SPC, HAILCAST can be run manually using observed or forecast soundings as part of the National Centers Advanced Weather Interactive Processing System Skew T Hodograph Analysis and Research Program (NSHARP; Hart et al. 2003). HAILCAST can utilize gridded NWP model data to produce plan-view hail-size forecast graphics. Verification of HAILCAST has not been done using the forecasts from gridded model data.

In this study, the model was tested using a large database of individual observed hail proximity soundings. A brief description of the hail model is provided in section 2. The report database is discussed in section 3, which is followed by a description of the methodology used to verify the model forecasts in section 4. Finally, some test results are presented.

2. The HAILCAST model

HAILCAST is a one-dimensional coupled cloud and hail growth model, developed initially by Poolman (1992) and then improved upon by Brimelow et al. (2002). Using an atmospheric sounding as input, the model produces an ensemble of updrafts based on systematic perturbations of the control surface temperature and dewpoint. At the forecaster's discretion, the control temperature and dewpoint can be modified to create a more representative sounding. A set of 25 ensemble members is produced by varying both the temperature and dewpoint by 1°C from their control values in increments of 0.5°C . Each of the 25 surface parcels are lifted to ascertain if they can reach their LFC. Those that reach the LFC are considered to produce deep convection, and these convective members are evaluated for hail size potential. This aspect of the hail model will be described in more detail in section 5.

HAILCAST utilizes the energy shear index (ESI; Brimelow et al. 2002), which is the product of the surface-based CAPE and the 850 mb–6 km AGL bulk wind shear. The purpose of the ESI is "to account for the combined effects of buoyancy and vertical wind shear on the updraft duration, with larger products indicating an increased potential for long-lived updrafts" (Brimelow 1999). Depending upon the magnitude of the ESI, varying degrees of lateral and cloud-top entrainment are applied to the updraft. HAILCAST decreases the amount of entrainment as the instability and shear increase, resulting in maximized updraft speeds approaching the theoretical buoyancy-derived values associated with supercells (Bluestein et al. 1988). ESI regulates the updraft duration with a maximum lifetime of 60 min. After a $300\text{-}\mu\text{m}$ embryo is introduced at cloud base (calculated using a surface-based parcel), the hail model allows the embryo to ascend within the model-derived updraft and grow until either the updraft collapses or the hailstone reaches the surface. The embryo, injected at cloud base, is given an initial upward velocity of 4 m s^{-1} . However, it is important to note that surface-based convective inhibition (CIN) or a capping inversion may prevent parcels from reaching their LFC. Thus, if one or more members remain capped (e.g., they are unable to produce deep convection), no hail will be produced for those members. For a more detailed description of HAILCAST, see Brimelow et al. (2002).

3. Data and methodology

a. Hail report database

HAILCAST was evaluated in Alberta using a relatively high quality report database of hail size from a

Total Hail Reports 1989-2004

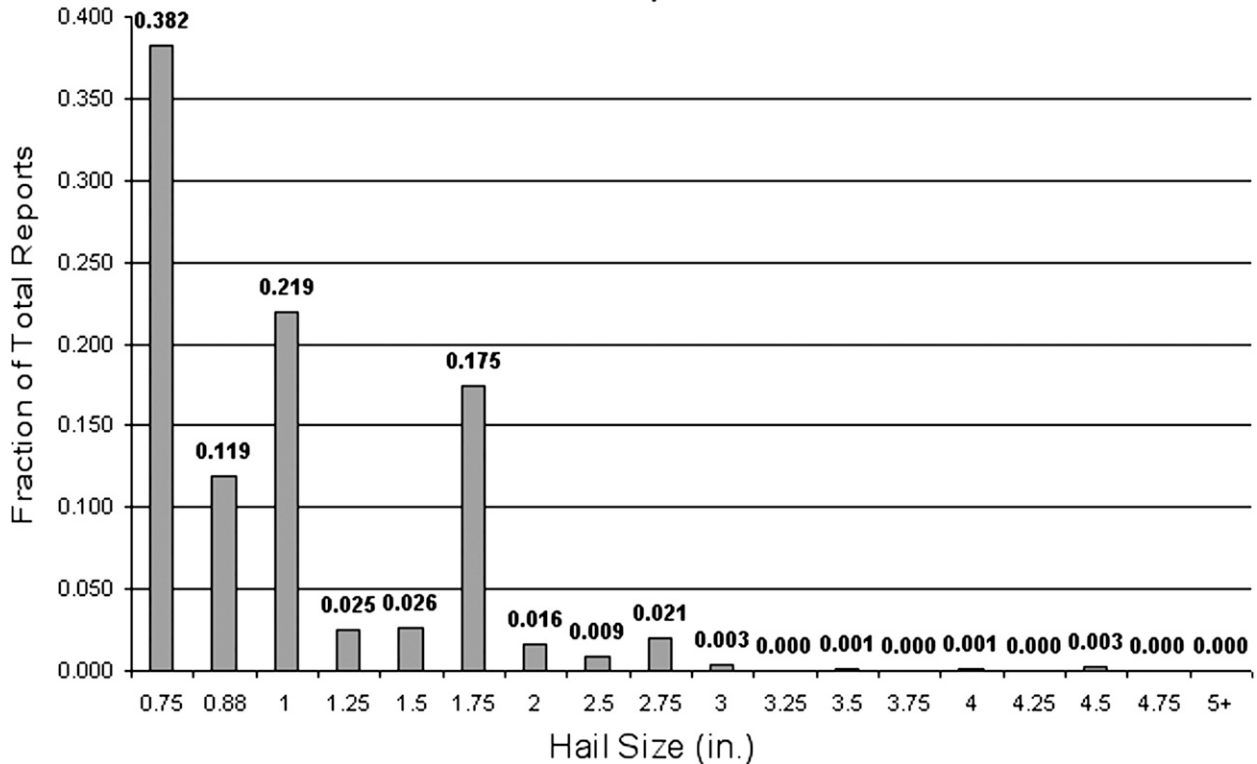


FIG. 1. Distribution of severe hail reports by size over the contiguous United States from 1989 to 2004. Note the apparent golf-ball-size anomaly at 1.75 in. (4.45 cm). There were 137 837 total reports for the period.

high-density hail observer network (Brimelow et al. 2002). The overall quality of the severe report database (*Storm Data*) in the United States is much less reliable, with the vast majority of the hail sizes being estimated, and only recently has information about whether a stone was measured or estimated been included in the local storm reports. In addition, the largest hail size reported is not necessarily the largest stone produced by a storm (Amburn and Wolf 1997), although growing spotter networks are helping to mitigate this factor. Therefore, even if a perfect hail size forecasting tool existed, the quantitative verification of hail size forecasts based on the current U.S. hail database would still be challenging. Despite these limitations, this is the only long-term database that we have at our disposal for the purpose of verifying HAILCAST. Recently, the NSSL Severe Hail Verification Experiment (SHAVE) was conducted and employed a high-resolution verification database, the benefits of which are discussed in Smith et al. (2007).

The most common reported diameter is 0.75 in. (1.9 cm), the threshold for severe hail as defined by the National Weather Service (NWS), averaging around 4000 reports per year. The combination of 0.75-, 0.88- (2.2 cm), and 1.00-in. (2.5 cm) hail accounts for about

6200 reports per year over the 16-yr period, or 72% of all reports. In contrast, an average of only 183 reports of baseball-size hail ($D = 2.75$ in. or 7.0 cm) were received each year, or about 2% of the total reports.

Analysis of the U.S. severe hail database indicates that very large hail is observed much less frequently than smaller hail, indicating an approximate inverse relationship between hail size and the frequency of occurrence. Looking at Fig. 1, this is generally the case, except that the number of golf-ball-size hail reports in relation to adjacent sizes appears anomalously high, especially given the mass differential between a spherical 1.75-in. (4.45 cm) stone and stones in the smaller size categories. For the period 1989–2004, 1.75-in. (“golf-ball size”) hail was reported almost as frequently as 1.00-in. (“quarter size”) hailstones, and much more frequently than 0.88-in. (“nickel size”) hail. This is not a new phenomenon. For the period 1955–2002, reports of golf-ball-size hail actually exceeded those of 1.00-in. hail and were the second most frequent size reported (Schaefer et al. 2004). An informal study on the accuracy of hail-size reporting (Baumgardt 2008) showed that spotters are much less likely to report 1.25-in. (3.18 cm) to 1.50-in. (3.81 cm) hail sizes than they are to report 1.00-in.

(quarter size) or 1.75-in. (golf ball) size hail. This is consistent with the results in Fig. 1, which shows a relatively low number of 1.25- and 1.50-in. reports. Further, personal communication with longtime Oklahoma residents living near the world-wide climatological maximum of giant hail frequency (Doswell et al. 2005) reveal that golf-ball-size hail is observed much less frequently compared to the occurrence of quarter-size or smaller hail.

There is also reason to question the 4.50-in. (11.4 cm) hail reports, the default size used by the NWS to represent softball-size hail. The first-ever softball tournament at the 1933 World's Fair in Chicago used a 4.50-in.-diameter softball. However, the most common men's league softball in use today has a diameter of 3.80 in. or 9.65 cm (women's league softballs measure 3.50 in. or 8.89 cm). Consequently, observers are likely to associate the classification of "softball sized" hail with a hailstone of 3.80-in. diameter, and not 4.50 in. As a result, it is highly likely that many of the allocations of 4.50 in. to reports of softball-size hail in the database are gross overestimations of the true size of the hail. This confusion has important implications, not only for verification purposes, but also for inferring the associated thunderstorm intensity. For example, a 4.50-in.-diameter stone has 63% more mass when compared to a 3.80-in. (true softball size) stone. This is not insignificant, as the updraft required to produce a 4.50-in. stone typically must be significantly stronger than that required to support a 3.80-in. hailstone. To put this into perspective, the 1970 Coffeyville, Kansas, stone had a mass near 750 g, while a solid 4.50-in. sphere of ice, assuming an ice density close to 0.9 g cm^{-3} (Knight and Knight 2001), would have a mass of just over 700 g. It is unlikely that near-world-record stones fall with such frequency as the report database would indicate. For the purposes of this study, 4.50-in. (softball) hail reports will be assumed to be 3.65-in. (9.27 cm) spheres, an average of a modern men's and women's softball sizes.

b. Building the sounding database

A database of observed severe hail proximity soundings from the contiguous United States spanning nearly 15 yr (May 1989–April 2004) was constructed for the purpose of evaluating the HAILCAST model. It is important to note that only soundings supporting surface-based convection were included in the dataset, as the original version of HAILCAST was unable to simulate elevated convection [storms with inflow rooted above a surface-based stable layer; Corfidi et al. (2008)]. Although HAILCAST has since been modified to simulate elevated convection, estimating representative parcel

conditions for elevated sounding cases (especially historical ones) is extremely problematic. This is because the mesoscale variability of conditions aloft is not well sampled, whereas surface data contain substantially greater time and space resolutions to sample surface-based convective parcel characteristics. Thus, soundings with only elevated instability were not included in the database.

The SPC software package SVRPLOT (Hart 1993) was used to sort and plot hail reports from *Storm Data*. Hail reports were only included if they were observed within 100 n mi (185 km) of an upper-air radiosonde site, and occurred between 2100 and 0200 UTC (2.5 h from 2330 UTC). Reasons for using 2330 UTC as the representative sounding release time include the fact that it takes time for the balloon (typically released shortly after 2300 UTC) to reach the mid- and upper levels of the atmosphere, and the possibility of late or multiple releases. These proximity criteria are similar to those used by Craven and Brooks (2004), but are also rather arbitrary in nature (Brooks et al. 1994). Given the fact that most hail events occurring near 1200 UTC are elevated in nature due to temperature inversions in the boundary layer (and are thus unrepresentative surface temperature measurements), only afternoon cases were considered in this study. This is also consistent with Brimelow et al. (2002), who used only afternoon hail cases. For an in-depth discussion of the challenges associated with choosing proximity sounding criteria, see Brooks et al. (1994).

Hail reports were categorized as either SIG (≥ 2.0 in.) or NON-SIG (< 2.0 in.) according to the size criteria specified in section 1. On days with multiple hail reports within 100 n mi of a sounding site, the largest hail report within that distance from the sounding site was recorded. Care was taken when choosing the NON-SIG events, such that there were no SIG hail reports within 300 n mi of the NON-SIG reports. However, the majority of NON-SIG hail events were on days when no SIG hail was reported anywhere within the contiguous United States. Soundings were excluded if they were contaminated by convection, which typically included deep saturated profiles, unusual temperature lapse rates, suspect winds (e.g., a cold front passage), or missing data (Fig. 2). Archived hourly surface observations were used to support whether or not the sounding was released in the same low-level air mass that was ingested by the hailstorm's updraft, in order to obtain a representative proximity temperature and dewpoint.

The tested version of HAILCAST does not use winds below 850 hPa, so erroneous or unrepresentative winds below this level are irrelevant to the model's calculations. In the few cases where the surface pressure was just below 850 hPa, the surface wind was used in place

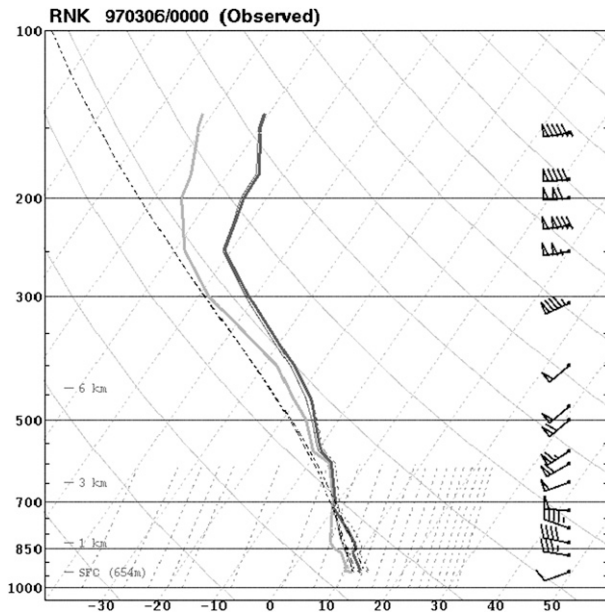


FIG. 2. Example of a convectively contaminated sounding at Roanoke, VA, at 0000 UTC in March 1997. The degree of saturation and lapse rate profile suggest that a balloon likely entered a convective cloud.

of the 850-hPa wind. Because the model uses surface observations to reconstruct a well-mixed boundary layer, rain-cooled air below the LCL is also irrelevant if surface conditions prior to contamination exist for sounding modification. However, for the purposes of this study, soundings with contaminated low levels were not used, in order to minimize the likelihood that other aspects of the sounding were also contaminated.

After applying the selection criteria, a dataset of 942 soundings was constructed. Only 28 of these soundings (3%) were capped for a surface-based parcel after it was modified for surface conditions (methodology explained in section 5). Given that the model is unable to produce deep moist convection on a capped sounding using a surface-based parcel, these soundings could not be used; thus, 914 severe hail proximity soundings remained (Fig. 3). Of these, 490 were associated with SIG hail and 424 with NON-SIG hail. The largest number of hail proximity soundings is over the plains region, consistent with the climatological occurrence of severe and significant hail (Doswell et al. 2005).

c. Selecting the control surface temperature and dewpoint for model input

As with previous hail forecasting techniques, the maximum forecast hail size is highly dependent on the updraft properties and, thus, the parcel properties used to calcu-

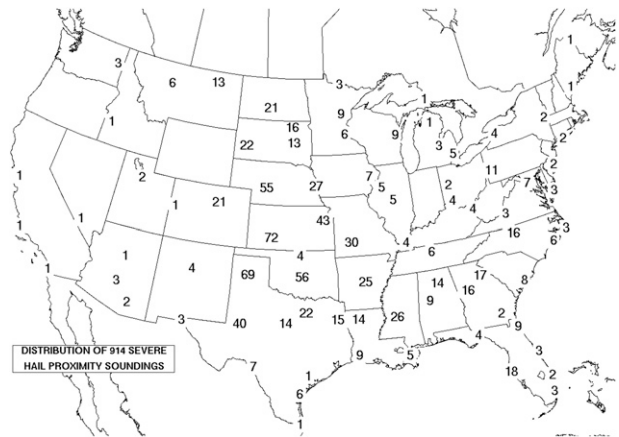


FIG. 3. Location and number of severe hail proximity soundings used during 1989–2004. The database includes soundings in every month.

late the updraft strength. Therefore, it is critical to employ a robust method for selecting the maximum representative surface temperature and dewpoint values in the inflow sector that will characterize the most intense updraft.

Selecting a representative control surface temperature and dewpoint for each proximity sounding is difficult and necessarily subjective given the relatively sparse observational data. Although one simply could use the observed surface conditions at the sounding site at 0000 UTC, the possibility of unrepresentative near-storm surface conditions would be increased, especially if the storm is located some distance from the sounding site. Because it is impossible, especially for such a large database, to accurately sample the exact properties of the air being ingested by each hailstorm, one approach is to estimate an upper limit of the temperature and dewpoint, which will provide an upper limit to the surface-based CAPE. Given that the majority of surface based hailstorms are observed at the time of day when the near-surface lapse rate is well mixed and perhaps super-adiabatic, it is much more likely that the highest surface potential temperature observed between 2100 and 0200 UTC is an overestimate of the actual average potential temperature between the surface and cloud base, rather than an underestimate. Overly high temperatures can result if the surface measurement is taken near concrete buildings or asphalt runways. Therefore, using the warmest surface temperature found will define a reasonable upper limit. The same logic applies to determining the dewpoint. Dewpoints typically do not increase with height in the subcloud layer, so the highest dewpoint found at the surface is often an overestimate of the average dewpoint below cloud base (Craven et al. 2002).

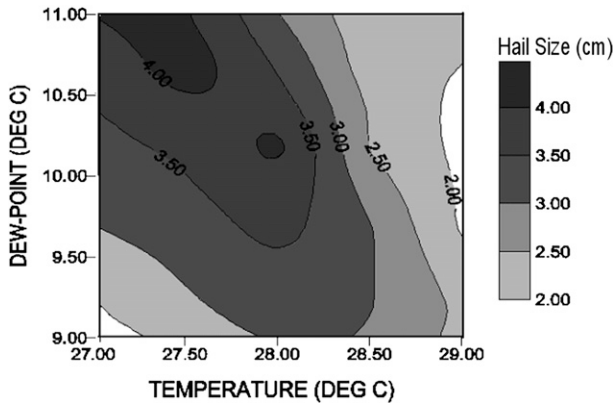


FIG. 4. Graphical example of forecast hail size for a HAILCAST simulation showing hail-size variability as a function of temperature and dewpoint.

Using available surface observations, the following guidelines were employed to find a representative upper limit for the temperature and dewpoint to be used:

- The highest temperature (T_{max}) and dewpoint (T_{dmax}) observed between the hail report time and 2.5 h prior to the report time were recorded, within the inflow air mass ahead of the storm.
- These maximum temperature and dewpoint values did not have to occur at the same observation point, but they had to be within the same surface air mass feeding the storm.

- Observations upwind of the hail report (with respect to the surface flow) were preferred. An exception to this guideline was made if some of the observations were contaminated by precipitation, or if surface winds were light and variable.

HAILCAST takes the input (control) surface temperature and dewpoint, and then varies each of them by 1°C in 0.5°C increments to produce an ensemble forecast of 25 surface parcels. The perturbation method assumes incomplete sampling of the surface conditions, such that variability on the mesoscale exists and needs to be accounted for. A graphical example is shown in Fig. 4. Using the case example shown in Fig. 5, the highest temperature–dewpoint combination observed would be 92°–73°F (33°–23°C). Thus, the final control values input into the model would be 90°–71°F (32°–22°C) (the high values minus 1°C or about 2°F). The model would then produce an ensemble of 25 parcels, ranging from 88°–69°F to 92°–73°F (31°–21°C to 33°–23°C). If one were to enter the maximum observed temperature and dewpoint directly into the model, overly buoyant ensemble members would result, with the most unstable parcel in this example being 94°–75°F (34°–24°C). A parcel of that magnitude is not supported by the available surface data.

This method of selecting the temperature and dewpoint along with the HAILCAST ensemble perturbation process also helps account for the effects of vertical

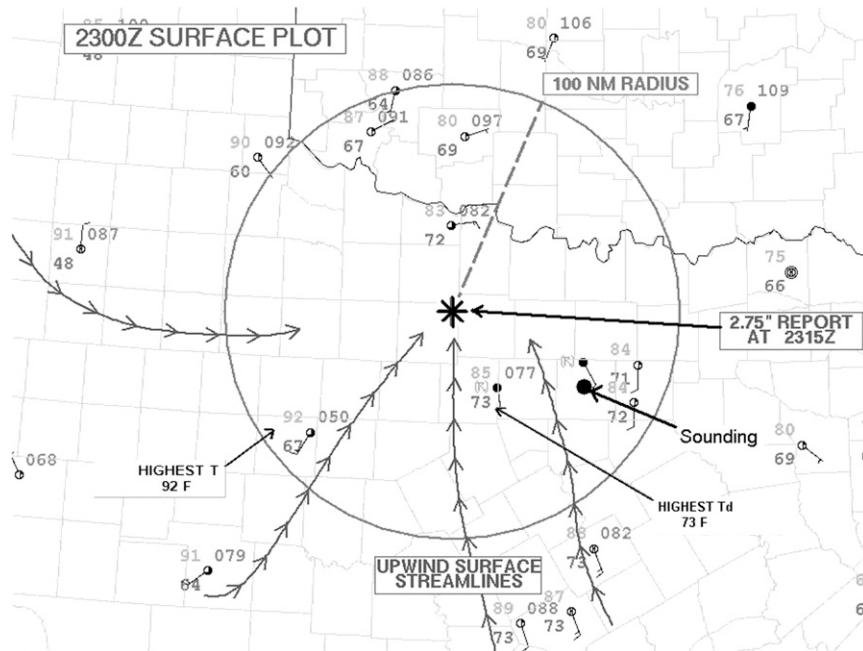


FIG. 5. Graphic illustrating the method used for obtaining T_{max} and T_{dmax} . Highest values in this case would be 92°F (33°C) and 73°F (23°C), respectively. Control values for the hail model would then be $T = 90°F$ (32°C) and $T_d = 71°F$ (22°C).

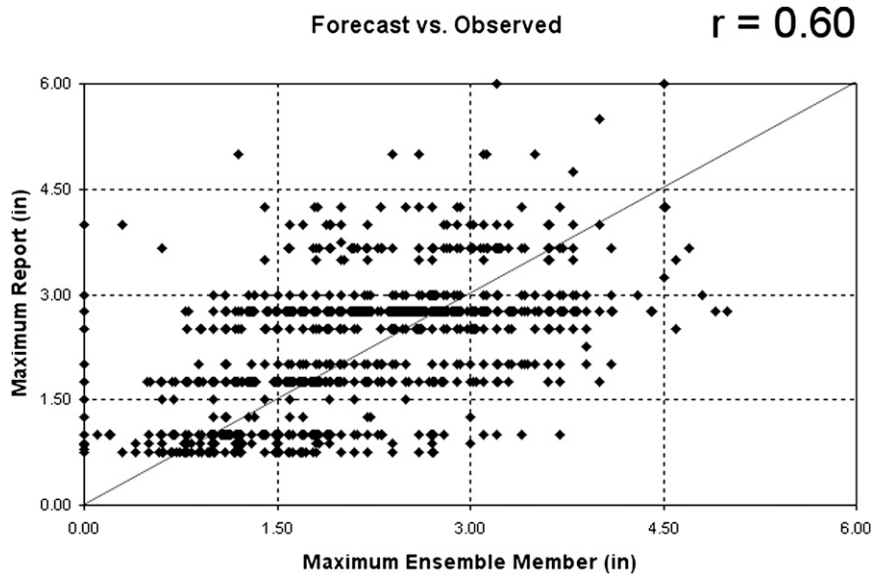


FIG. 6. Scatterplot of observed hail diameter vs ensemble maximum diameter. The diagonal line represents a perfect forecast. Some plots are overlapping.

mixing and inhomogeneous temperature and dewpoint fields in the vicinity of the storms. It has been shown by Craven et al. (2002) that, for the central United States, using a 100-mb mixed layer temperature and dewpoint when computing the updraft properties appears to be more appropriate than using a surface-based parcel, which tends to result in excessively warm and moist lifted parcels. This is especially true when dealing with skin layers of moisture that can be unrepresentative of convective potential, and that result in overestimates of the actual buoyancy. Therefore, using $T_{\max} - 1^{\circ}\text{C}$ and $T_{d\max} - 1^{\circ}\text{C}$ as control (model input) values will account indirectly for some of the effects of mixing.

There are many methods that could be used to estimate a “representative” near-storm environment, but all require assumptions about the horizontal distributions of temperature and moisture. Therefore, it is difficult to assess which method would be best. Using Fig. 5 as an example, one *could* assume that the surface dewpoint would be less than 71°F and the surface temperature less than 90°F at the location of the storm report. This *may* be a good assumption, but it could also be said that assuming a linear gradient is an oversimplification. Conversely, it could be argued that there is a *higher* dewpoint and temperature at the storm report location if you do not assume a linear moisture gradient. One could simply use the single closest surface observation to the storm report location, but there is an increased risk of contamination by convection or a bad observation. For these reasons, the authors considered multiple surface observations in the near-storm environment.

4. HAILCAST performance

a. Hail category forecasts

It is evident from Figs. 6 and 7 that the maximum ensemble member exhibits a positive association with the observed hail size, with the forecast values centered approximately around the perfect forecast line, while the ensemble mean differs by having a strong tendency to underestimate the maximum observed hail size. It is important to note that the ensemble mean only includes members whose parcels reach the LFC.

The linear correlation coefficient r for the maximum ensemble member is 0.60 (Fig. 6), versus 0.61 (Fig. 7) for the ensemble mean. Evident from Fig. 6 is that the model fails to produce hail for a subset of the soundings. Specifically, the model failed to produce hail in 18% of the proximity soundings. The vast majority of these (88%) were NON-SIG proximity soundings.

A common type of sounding profile for which the model fails to produce hail is found across the southeastern United States or the Gulf coast during the summer. Such environments are characterized by small vertical wind shear and weak midlevel lapse rates. Under such conditions, the updraft contains very high liquid water contents. Additionally, because of the relatively warm updraft temperatures found in such environments the hailstone will not be capable of freezing most of the intercepted supercooled water. Consequently, most of the (unfrozen) intercepted water will then be shed, thereby reducing the growth potential and increasing the chances that the stone will melt before reaching the

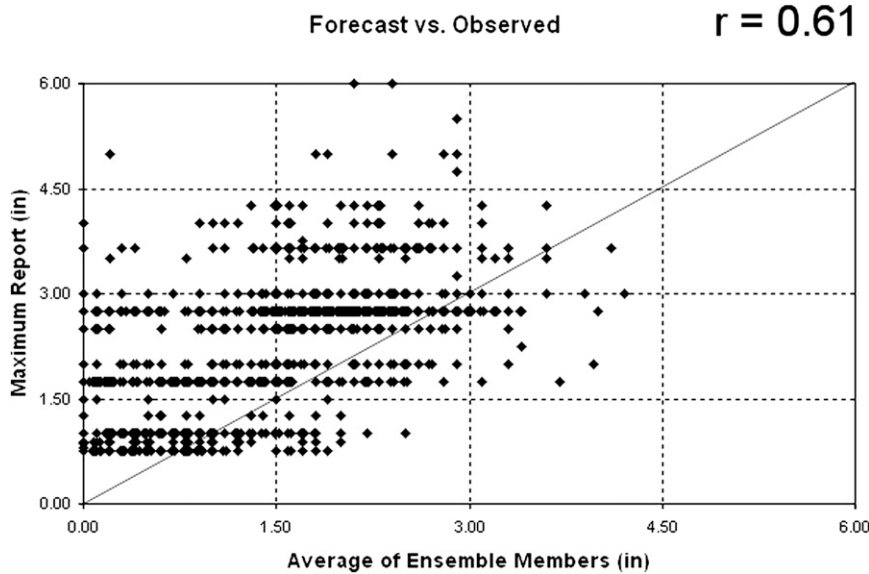


FIG. 7. As in Fig. 6 but for the observed hail diameter vs ensemble mean diameter. Some plots are overlapping.

ground. The growth potential is further limited by the relatively short updraft durations predicted by the model owing to low shear, and increased melting on account of the high FZL. A sample summertime sounding from the southeastern United States depicting a warm and moist vertical profile with little vertical shear is shown in Fig. 8. Despite having relatively large CAPE, these types of soundings are typically only capable of producing relatively small hail.

Figure 9 shows the average maximum member for all cases within various size groups, as well as the average ensemble mean for the same groups. The groups (and number of soundings in each) are 0.75 in. (107), 0.88 in. (38), 1.00 in.(109), 1.38 in. (the average size of all 1.25- and 1.50-in. reports, numbering 23), 1.75 in. (144), 2.00 in. (52), 2.50 in. (67), 2.75 in.(204), and 3.62 in. (9.19 cm, the average size of all reports ≥ 3.00 in. numbering 170). For example, there are 204 baseball-size (2.75 in.) soundings in the database. The average size forecast by the hail model for this group of soundings when using the maximum ensemble member is 2.41 in. (6.12 cm), while the average size forecast by the ensemble mean is only 1.76 in. (4.47 cm). All 1.25- and 1.50-in. reports were grouped given a relatively small sample size of each. Figure 9 shows that the average maximum ensemble member line corresponds closely to the perfect fit line, albeit with a negative bias. The average ensemble mean value parallels the maximum but with a greater negative bias.

Although the average ensemble maximum member has less bias than the average ensemble mean, using

a single maximum member may be less useful than using an ensemble mean. Figure 10 shows that the standard deviation of the average ensemble mean forecasts is smaller than that of the average maximum ensemble member for all size categories. A likely explanation for the larger variability when using the maximum ensemble

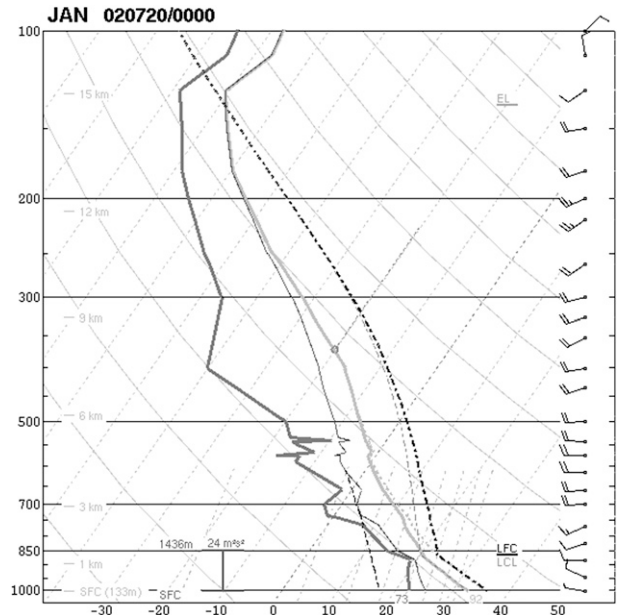


FIG. 8. Example of a summertime sounding exhibiting a warm and moist thermodynamic profile with weak midlevel lapse rates and weak wind shear. The sounding is from Jackson, MS, at 0000 UTC 20 Jul 2002.

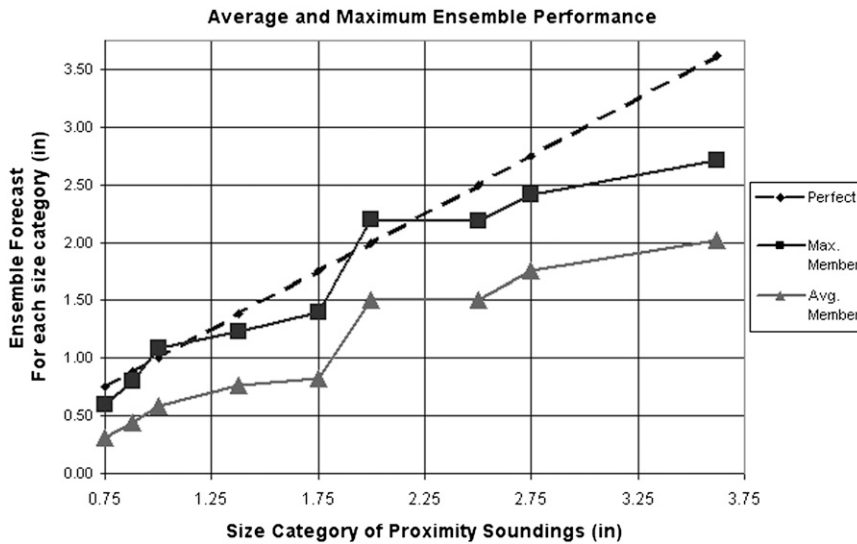


FIG. 9. Plot of the average of the ensemble maximum hail size (black line–squares) and average ensemble mean hail size (gray lines–triangles) of all forecasts for each size category. The dashed line indicates perfect reliability.

member is the use of a single combination of temperature and dewpoint for each sounding, which increases the likelihood of selecting an unrepresentative updraft parcel. Also, if the ESI of an ensemble member crosses a categorical threshold (Brimelow et al. 2002), it will result in updraft properties being changed. The mean of all 25 members appears to reduce these effects.

Figures 9 and 10 suggest that applying a bias correction to the ensemble mean will result in improved reliability and accuracy in the forecast of maximum hail size. The calculated mean bias across all sizes is

–0.77 in. (–2.0 cm). However, from Fig. 9 it is seen that much of the bias can be attributed to the larger mean forecast error associated with very large hail (generally ≥ 2.00 in). For hail ≤ 1.50 in. diameter, the bias is only –0.48 in. (–1.2 cm). Thus, applying a +0.77 in. correction equally for all situations would tend to overforecast the hail size, given that, climatologically, 72% of all reports in the database are ≤ 1.00 in. diameter. A practical approach may be to first determine the likelihood of very large versus small hail, and then to adjust the ensemble mean output accordingly to arrive at a more

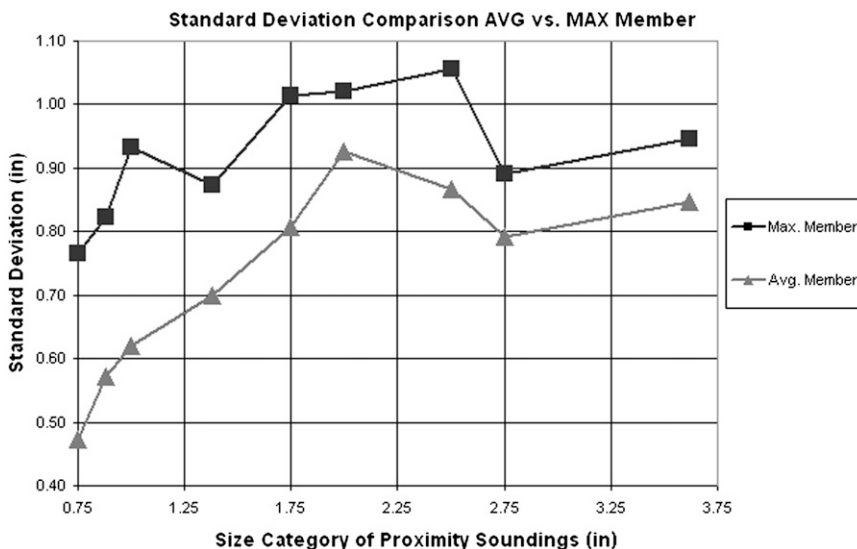


FIG. 10. As in Fig. 9 but for the standard deviation.

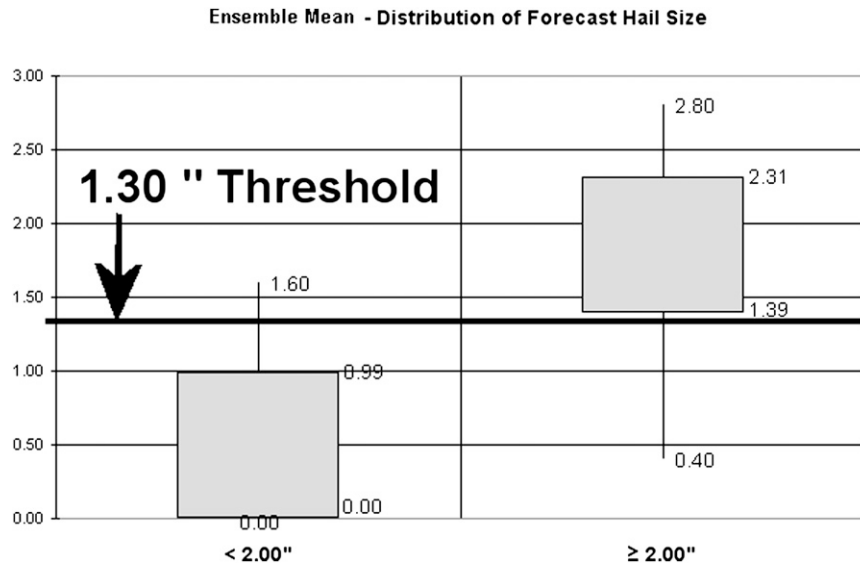


FIG. 11. Distribution of hail model output (ensemble average) for all soundings using two size categories. The box is the middle 50% of the distribution and the top (bottom) whisker extends to the 90th (10th) percentile.

accurate hail size forecast. This is explored in the next section.

b. SIG versus NON-SIG forecasts

To test the model's ability to discriminate between SIG and NON-SIG environments, the dataset was divided into two subsets: $D \geq 2.00$ in. and $0.75 \leq D \leq 1.75$ in. There is significant interquartile separation between the two groups shown in Fig. 11, indicating that HAILCAST possesses promising skill when discriminating between SIG and NON-SIG hail environments. The degree of separation is undoubtedly affected by the 18% rate of hail model failure (most failures occur in the < 2.00 in. category), where the 10th and 25th percentile values for hail forecasts for the < 2.00 in. category are both zero. Even so, if the model produces no hail, one can still say with confidence that SIG hail is unlikely.

Calculation of skill scores for discriminating between SIG and NON-SIG hail shows that the combination of the critical success index (CSI; Donaldson et al. 1975) and the true skill statistic (TSS; Doswell et al. 1990) is optimized when using an ensemble mean threshold value of about 1.30 in. (Fig. 12). An ensemble mean forecast above this threshold indicates a deterministic forecast of SIG hail, while values below indicate NON-SIG hail. Apparent from Fig. 11 is that the 1.30-in. threshold can be adjusted to more closely approximate reality, by adding 0.70 in. to the ensemble mean forecast. When this is done, the discrimination threshold is now 2.00 in., which divides SIG and NON-SIG hail by definition (Fig. 13).

Finally, filtering out hail reports that are near the threshold value of 2.00 in. should decrease the noise (e.g., hail-size uncertainty) and better reflect HAILCAST's ability to discriminate between SIG and NON-SIG events. A total of 718 cases remain after removing all cases ≥ 1.75 in. and < 2.50 in. Figure 14 shows increased interquartile separation between the two size groups, with the 25th percentile output for hail ≥ 2.50 in. coincident with the 90th percentile for ≤ 1.50 in. hail soundings. The probability of detection (POD) of SIG hail for the filtered dataset increases to 0.81 while the CSI and TSS values increase to 0.75 and 0.68, respectively (Table 1). This confirms that HAILCAST exhibits substantial skill in discriminating between environments favorable for hail residing at both ends of the severe hail spectrum.

The forecast of SIG or NON-SIG hail is thus most uncertain in the heart of the overlapping regions for the two size groups, centered approximately between 1.60 and 1.80 in. for the bias-corrected ensemble mean (Fig. 13). This is in part due to the hard threshold of 2.00 in. Interestingly, Fig. 10 shows that the standard deviation is maximized for ensemble mean forecasts of size groups near the 2.00-in. threshold, specifically from 2.00 to 2.50 in.

c. Possible explanations for the systematic negative bias

It was shown in Fig. 9 that the negative forecast bias for the ensemble mean increases with increasing hail size and is largest for hail ≥ 2.00 in. with bias values

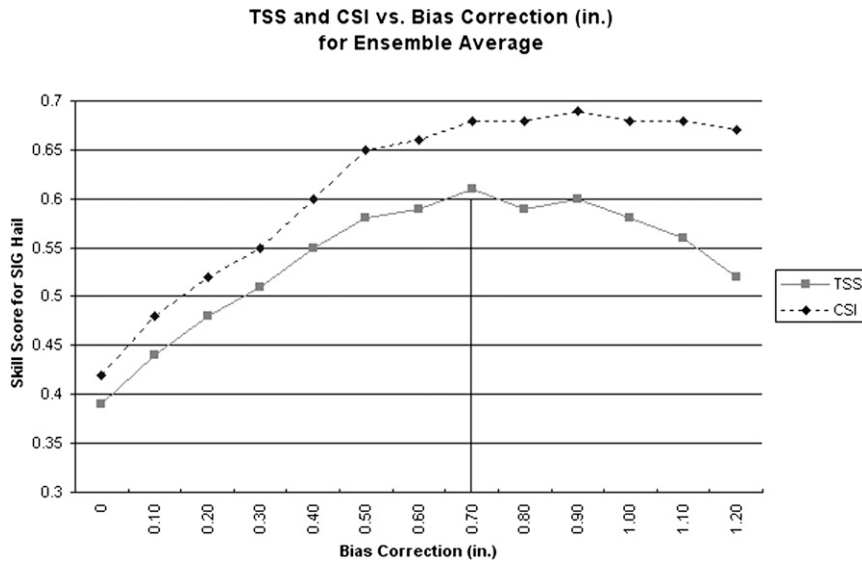


FIG. 12. TSS and CSI as a function of bias correction. TSS is maximized with a bias correction of 0.70 in.

>1.00 in. It is appropriate to examine this result in more detail to determine possible factors leading to the increasing low bias for larger hail sizes.

1) HAILSTONE MORPHOLOGY

One possible reason for the low bias is that HAILCAST assumes a perfect, spherically shaped hailstone, while in reality large hailstones are rarely perfect spheres. Instead, most hailstones are triaxial ellipsoids (García-García and List 1992). Barge and Isaac (1973) measured both the minimum and maximum dimensions of a large

number of hailstones from Alberta hailstorms. They found that the modal value of the aspect ratio (α ; minor axis length divided by major axis length) was between 0.75 and 0.79. Similarly, Matson and Huggins (1980) determined that the mean aspect ratio of their hailstone sample was 0.77. There is evidence that large hailstones become more oblate as they grow. For example, Browning and Beimers (1967) examined 90 large stones and concluded that the stones became more oblate (decreasing aspect ratio) as the maximum hail dimension increased.

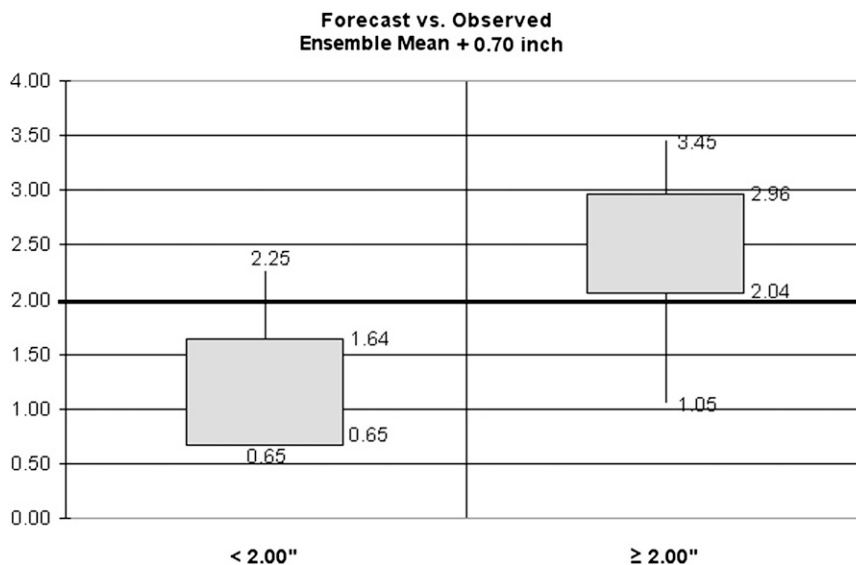


FIG. 13. As in Fig. 11 but with a 0.70-in. (1.8 cm) bias correction applied.

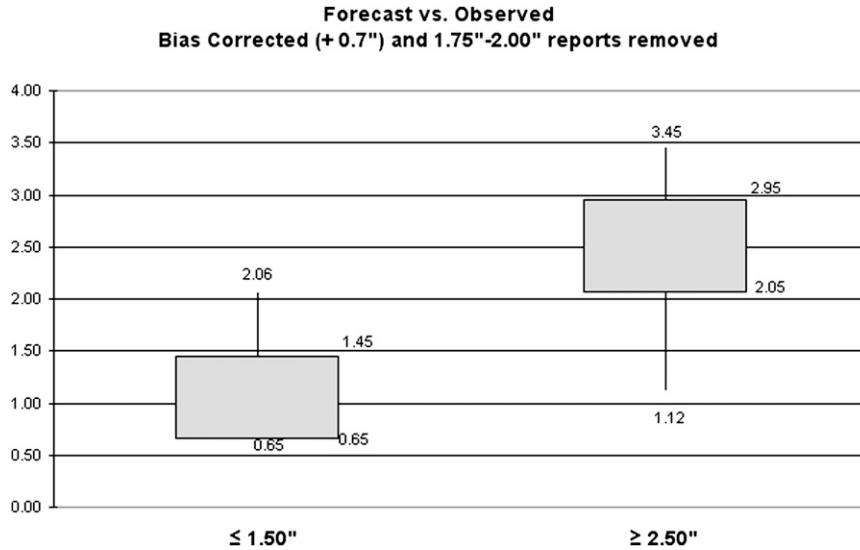


FIG. 14. As in Fig. 13 but with all reports ≥ 1.75 in. (4.35 cm) and < 2.50 in. (6.35 cm) removed. Separation between the two size categories increases as a result of this filtering.

The size of an oblate hailstone is often quantified through a parameter known as the equivalent diameter, D_{eq} . The equivalent diameter is the diameter that an oblate hailstone of a given mass would have if it were perfectly spherical. Here, D_{eq} is a function of an oblate hailstone’s aspect ratio α and major diameter D :

$$D_{eq} = \alpha^{1/3} D. \tag{1}$$

The equivalent diameter is a useful parameter when comparing the sizes of spherical and oblate hailstones, because it is directly related to hailstone mass M :

$$M = \frac{\pi}{6} \rho_H D_{eq}^3, \tag{2}$$

where ρ_H is the hailstone density. Applying an aspect ratio of 0.75 to Eq. (1) yields an equivalent diameter that is about 10% smaller than that of the length of the major axis. For a highly oblate stone having an aspect ratio of 0.50, the equivalent diameter would be some 20% smaller than the length of the major axis. Consequently, given that observers are likely to report the largest dimension of the hailstone (major axis), it is expected that estimations of the true hail mass will be overestimated when using reports from observers, especially when stones have uneven surfaces, or if they are very oblate.

For example, an image of giant hailstones from 28 April 2002 is shown in Fig. 15. These stones may be the ones that were reported as 4.50-in. “softball size” stones in *Storm Data*. A 3.00-in. (7.62 cm) hailstone was reported at nearly the same time and location as the 4.50-in. report. Although the major axis of the larger

stone was about 4.50 in., the total mass of the stone was likely much less than that of a solid 4.50-in. sphere. Specifically, the aspect ratio of the larger stone was approximately 0.62, which results in an equivalent diameter of 3.85 in. (9.78 cm), so one would expect this stone to have a mass closer to 444 g and not 705 g as predicted using a spherical diameter of 4.50 in. On this day, HAILCAST predicted a maximum hail diameter of 2.80 in. (7.11 cm) using a modified 0000 UTC proximity sounding.

2) HAILSTONE TRAJECTORIES IN HAILCAST

Another possible explanation of the systematic low bias in hail size can be attributed to a limitation of the 1D nature of the model to simulate realistic trajectories through the storm. Specifically, in very high CAPE environments, the cloud droplet introduced at cloud base ascends rapidly through the depth of the modeled storm. In such situations, the hydrometeor has very little time to grow during its rapid ascent before it enters the glaciated zone (temperature $< -40^\circ\text{C}$). The absence of supercooled water and the low collection efficiency of ice particles then preclude rapid growth of the stone.

TABLE 1. SIG hail contingency table and skill scores for the ensemble average with reports both ≥ 1.75 in. (4.45 cm) and < 2.50 in. (6.35 cm) removed.

HIT	MISS	FA	CN
354	84	35	245
POD	FAR	CSI	TSS
0.81	0.09	0.75	0.68

CN = correct null; FA = false alarm.



FIG. 15. Large hailstones from the La Plata, MD, tornado event on 28 Apr 2002. Long axis of left stone is about 4.50 in. (11.43 cm). Compact disc diameter is near 4.70 in. (11.94 cm).

In exceptional circumstances, it may be possible for the stone to grow large enough to descend below the -40°C level and enter the mixed-phase region of the cloud where it can grow more rapidly before the updraft collapses. Alternatively, the stone remains suspended above the -40°C level until the updraft collapses. In the latter case, the hailstone then falls rapidly through the cloud because there is no updraft to prolong its residence time in the hail growth zone, and the stone is likely to be smaller than would otherwise be suggested by the environmental conditions. Unfortunately, this limitation of the model is difficult to overcome without invoking a 2D or 3D cloud model.

3) HAILCAST RELATIONSHIP TO CAPE

A common method used to estimate maximum hail size is to calculate the maximum theoretical updraft strength due to buoyancy alone and then calculate what size of hailstone can be supported by this velocity. This simple parcel theory method is a very inaccurate way of forecasting both the updraft velocity and maximum hail size (Doswell and Markowski 2004) for several reasons:

- Updraft strength is affected by water loading and entrainment, as well as wind shear in the near-storm environment, and pure parcel theory does not incorporate these factors.
- The maximum updraft velocity is at the equilibrium level (EL), which is typically at a much higher height than

where the most significant hail growth takes place. Supercooled water droplets are unlikely to exist at the EL for most storms, the exception being low-topped convective environments where CAPE values are smaller.

- Storm mode and updraft longevity are not taken into account.
- Microphysical processes relevant for hail formation are not considered.
- Dynamic vertical pressure gradient forces can contribute significantly to the updraft speed (Weisman and Klemp 1982, 1984; Rotunno and Klemp 1985; McCaul and Weisman 2001).

Thus, using the parcel theory method, large CAPE will predict large hail and result in very high false alarm rates (FARs; see Fig. 8 as an example). Figure 16 shows surface-based CAPE distributions for the filtered NON-SIG and SIG subsets. Note that there is much overlap between the CAPE associated with NON-SIG and SIG hail events. Comparing Fig. 16 with Fig. 14, the increased ability of HAILCAST to discriminate between SIG and NON-SIG hail events in strongly unstable environments becomes apparent when compared to using CAPE alone. These results are consistent with those of Edwards and Thompson (1998).

In addition, stronger updrafts do not necessarily correspond to larger hailstones. As can be seen in Fig. 4, the maximum hail size for this example actually *increases* with *decreasing* parcel buoyancy.

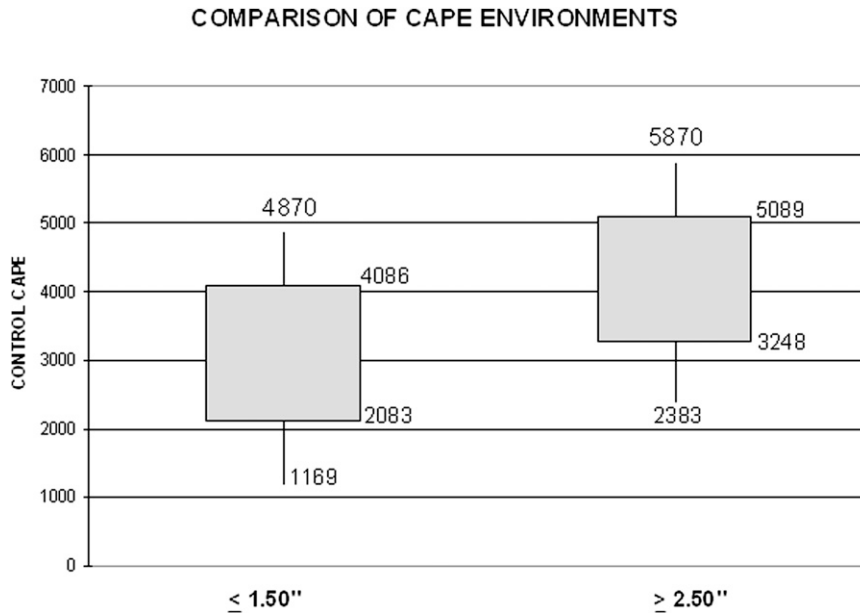


FIG. 16. Box and whisker plots comparing between the control surface-based (SB) CAPE environments for two size categories.

One of the few hail-forecasting methods available to NWS field forecasters is through an Advanced Weather Interactive Processing System (AWIPS) sounding algorithm. This is an overly simplistic method based mainly on the theoretical maximum updraft from CAPE. An excerpt from the AWIPS D2D users manual (NOAA 2008) states:

Maximum Hailsize. “The Maximum Hailsize represents the largest hailstone that can be supported by the undiluted parcel updraft (maximum vertical velocity). As a result, the size is exaggerated when compared to the size it is when it reaches the surface. This parameter is based on the equation given in BAMS, Vol. 62, No. 11, November 1981.”

An example of output from this algorithm is shown in Fig. 17. Note how the forecast maximum hail size using the AWIPS algorithm for this particular sounding is nearly 10 in. (25 cm) in diameter, when in fact the observed hail size was baseball-size hail (2.75 in.). Thus, the AWIPS algorithm does not appear to be a useful means of forecasting the maximum hail size based on environmental conditions, as any sounding with moderate to high CAPE will result in a giant hail size. When HAILCAST is run using the same sounding, the maximum forecast size is 2.20 in. (5.59 cm), with an ensemble mean of 1.70 in. (4.32 cm). In this case, HAILCAST clearly provided a much more realistic and superior hail-size forecast.

Figure 18 depicts the observed proximity sounding for the 22 June 2003 Aurora, Nebraska, record hail event

[7-in. diameter or 17.8 cm; see Guyer and Ewald (2004)]. The maximum forecast size using the modified Omaha, Nebraska (KOAX), sounding is 4.10 in. (10.4 cm), with an ensemble mean of 3.60 in. (9.1 cm). To place this hail forecast in perspective, of the 942 soundings in the database, only 10 (1%) had an ensemble mean of 3.60 in. or larger. In contrast, the AWIPS algorithm predicted a maximum hail size of 86 in. (218 cm).

5. Operational use of HAILCAST

It has been shown that bias-corrected ensemble mean output from HAILCAST provides a relatively reliable and accurate forecast of maximum hail size over the conterminous United States (CONUS). These results are based on a broad spectrum of storm types and environments. However, users of this model will have the most success by understanding how the model works, testing it, and carefully analyzing output. An experienced user will learn when to accept or reject the hail model output and, if necessary, make modifications to the input sounding.

Consider the following hypothetical example. Assume a particular forecast sounding gives widely varying hail size forecasts within a single 25-ensemble member run. The less unstable members, in combination with the deep-layer shear vector, result in shorter-lived convection, while the most unstable members may be enough to tip the storm mode into the supercell category, thus modifying the updraft properties and increasing the hail

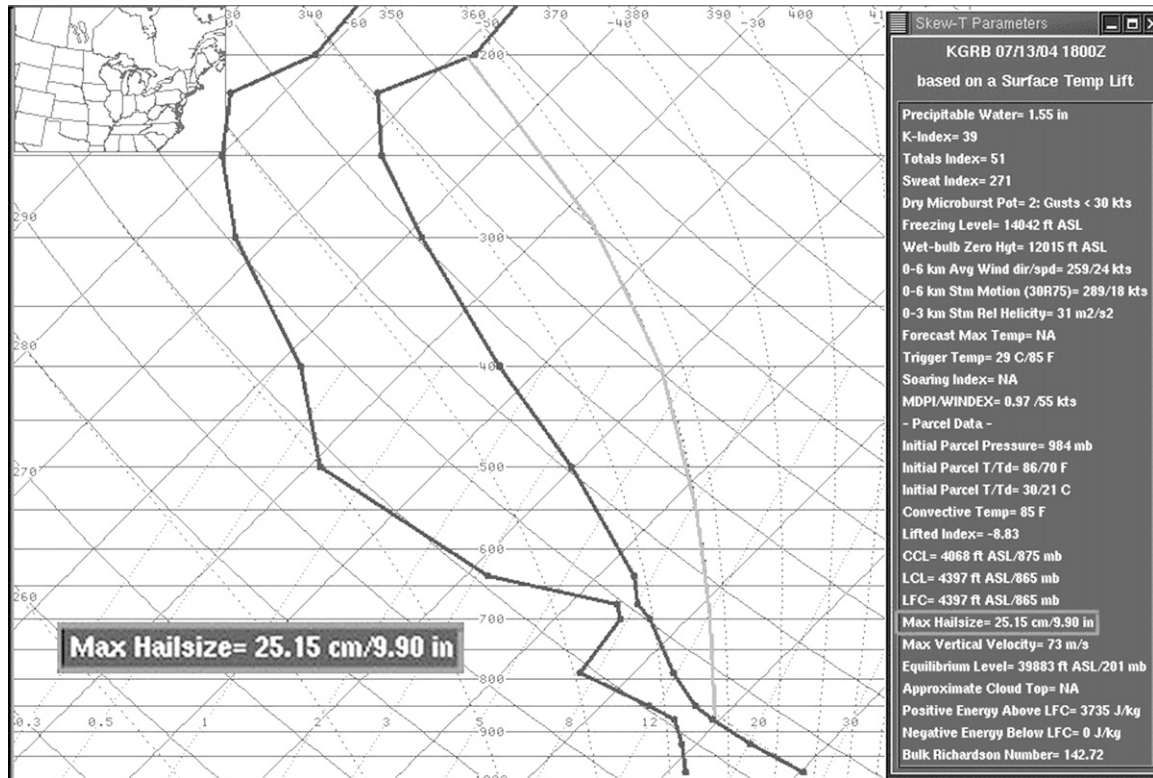


FIG. 17. Screenshot of the AWIPS sounding interface for 1800 UTC 13 Jul 2004 sounding at Green Bay, WI. Note that the “max hailsize” highlighted on the right is nearly 10 in. (25.4 cm).

size dramatically. In this case, the mean of the 25 members may not be very useful. The forecaster would have to use his or her judgment to determine which storm mode and ensemble members are most probable.

In some cases, the model fails to produce hail. Either the parcel and embryo initiated at cloud base do not rise past the LFC (there were zero “convecting members”) or the convecting members do not produce hail. Using an interactive sounding program, the forecaster may choose to modify the sounding by increasing the surface temperature sufficiently to break the cap, and/or by cooling the capping layer. This will increase the vertical velocity of the embryos and allow the parcel to rise past the LFC. A forecaster should be aware that use of these sounding adjustment methods, depending upon the magnitude of the changes, may introduce errors in hail size. Artificially increasing the surface temperature will have much less of an impact on the CAPE (Crook 1996) and the final hail size than would modifying the dewpoint, which is not recommended unless the forecaster feels the dewpoint is unrepresentative. This general approach of modifying a sounding permits a forecaster to conduct sensitivity tests of HAILCAST. However, forecasters should also be cautioned that having the capability

of modifying soundings does not mean the adjustments will be correct.

6. Potential HAILCAST improvement

In some cases, the lifted parcels cannot reach the LFC because of capping, and the model members fail to create a storm. While this typically means storms are unlikely to form, it is observed that supercells commonly persist into areas that are capped (Davies 2004). In these cases, the forecaster is left with a HAILCAST prediction of no hail. Modifying the code to introduce embryos at the LFC instead of the LCL may circumvent this problem. Similar work has been done on the model that allows it to run on elevated soundings in which the most unstable parcels are rooted above a surface stable layer, although this aspect of the model has yet to be evaluated.

In some cases, the ascending embryo may slow down significantly as it attempts to rise through a relatively stable capping layer. If the embryo fails to reach the LFC, then the simulation will simply die. However, the embryo can slow down so much that it grows substantially larger before accelerating above the LFC. This

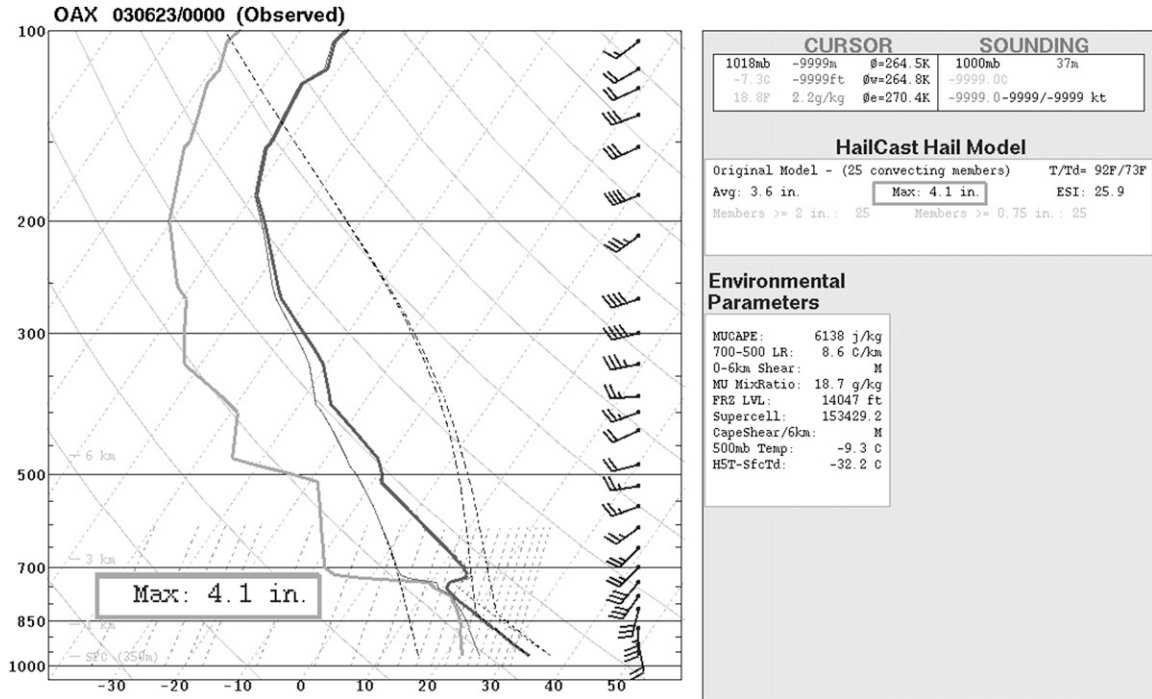


FIG. 18. NSHARP display at 0000 UTC from a KOAX sounding for the 22 Jun 2003 Aurora, NE, record hail event. The HAILCAST maximum forecast size is 4.1-in. (10.4 cm) diameter.

may result in a larger maximum hail size for the less buoyant ensemble members, with *smaller* hail sizes being produced for the more buoyant ensemble members using the same sounding.

The shear used in the model currently is the 850 hPa–6 km AGL bulk shear vector magnitude. The lower level of 850 hPa was originally chosen to approximate the surface heights found in the South African Highveld where elevations are near 1500 m MSL. The model may exhibit improved performance if the lower bound for the shear is changed to a surface- or a storm-relative inflow-layer average (Thompson et al. 2007). Also, giving the forecaster the option to force the storm mode and longevity would be helpful in situations where the model incorrectly diagnoses the storm type, and would allow examination of a wider range of possibilities. A version of the model that uses the effective bulk shear (Thompson et al. 2007) is currently being tested. A more complicated improvement would be to include the effects of wind shear on the updraft and contributions to vertical motion by way of dynamic vertical pressure gradient forces.

HAILCAST often fails to produce hail with very warm and moist soundings containing weak midlevel lapse rates and wind shear. A typical 700–500-hPa lapse rate in these cases would be $\leq 6.5^{\circ}\text{C km}^{-1}$, with surface–6-km bulk shear $\leq 10 \text{ m s}^{-1}$. Approximately 15% of the

soundings in the database fit this profile and are associated with severe hail (88% of those soundings are associated with hail ≤ 1.75 in. diameter). Additional testing should be done on these types of environments. Injecting the embryo at a higher level, such as the freezing level, may circumvent this problem, although further modifications to the model would be required to do this.

Other hail forecasting products that could potentially be useful would be to calculate the percentage of ensembles (i.e., probability) predicted to be greater than a specified threshold, and the standard deviation of the ensemble members (larger standard deviations would suggest greater uncertainty in the forecasts of specific hail size). Also, when producing hail maps using NWP prognostic soundings, Brimelow and Reuter (2009) found that the false alarm area for hail over central Alberta could be reduced by utilizing output from HAILCAST at only those grid cells where the Canadian Global Economy Model (GEM) was also predicting precipitation. However, rather than using deterministic forecasts of thunderstorms employing a single model, a more robust approach would be to couple the probabilistic output from HAILCAST with the model-based probability of thunderstorms, similar to those being produced using the Short-Range Ensemble Forecast (SREF) model (Bright et al. 2005).

7. Summary

HAILCAST is a practical tool for hail-size forecasting that can be run on observed or forecast soundings. Testing of HAILCAST using a large sample of proximity soundings reveals that the model can add great value to hail forecasts despite substantial deficiencies in the hail report database, significant errors associated with proximity soundings, and a relatively simplistic 1D model. The model is more physically realistic than traditional techniques as it grows an actual hailstone rather than estimating size based on ambient thermodynamic parameters alone. To its credit, HAILCAST does not forecast unreasonably large hail sizes in environments with high CAPE and/or vertical wind shear, and the output is relatively reliable.

HAILCAST appears to be the best tool presently available to forecast hail size. Although efforts are under way to test new improvements, the current version offers an objective hail size forecast that is scientifically sound and demonstrates considerable skill.

Acknowledgments. Thanks to Eugene Poolman and Julian Brimelow for providing HAILCAST to the SPC and also to Harold Brooks for setting up the international collaborative process. Thanks to John Hart, David Bright, and Rich Thompson of SPC for their programming support. Thanks also to Steve Weiss, Roger Edwards, Richard Thompson, and the anonymous reviewers for their suggestions.

REFERENCES

- Amburn, S., and P. Wolf, 1997: VIL density as a hail indicator. *Wea. Forecasting*, **12**, 473–478.
- Barge, B. L., and G. A. Isaac, 1973: The shape of Alberta hailstones. *J. Rech. Atmos.*, **7**, 11–20.
- Baumgardt, D., cited 2008: Hail estimation: How good are your spotters? [Available online at http://www.crh.noaa.gov/arx/hail_size_MSP.pdf.]
- Bluestein, H. B., 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.
- Bright, D. R., M. S. Wandishin, R. E. Jewell, and S. J. Weiss, 2005: A physically-based parameter for lightning prediction and its calibration in ensemble forecasts. Preprints, *Conf. on Meteorological Applications of Lightning Data*, San Diego CA, Amer. Meteor. Soc., 4.3. [Available online at <http://ams.confex.com/ams/pdfpapers/84173.pdf>.]
- Brimelow, J. C., 1999: Numerical modeling of hailstone growth in Alberta storms. M.S. thesis, Dept. of Earth and Atmospheric Sciences, University of Alberta, 153 pp.
- , and G. W. Reuter, 2009: Explicit forecasts of hail occurrence and expected hail size using the GEM–HAILCAST system with a rainfall filter. *Wea. Forecasting*, **24**, 935–945.
- , —, and E. R. Poolman, 2002: Modeling maximum hail size in Alberta thunderstorms. *Wea. Forecasting*, **17**, 1048–1062.
- Brooks, H. E., C. A. Doswell III, and J. Cooper, 1994: On the environments of tornadic and nontornadic mesocyclones. *Wea. Forecasting*, **9**, 606–618.
- Browning, K. A., and J. G. D. Beimers, 1967: The oblateness of large hailstones. *J. Appl. Meteor.*, **6**, 1075–1081.
- Changnon, S. A., 1972: Examples of economic losses, from hail in the United States. *J. Appl. Meteor.*, **11**, 1128–1137.
- Corfidi, S. F., S. J. Corfidi, and D. M. Schultz, 2008: Elevated convection and castellanus: Ambiguities, significance, and questions. *Wea. Forecasting*, **23**, 1280–1303.
- Craven, J. P., and H. E. Brooks, 2004: Baseline climatology of sounding derived parameters associated with deep, moist convection. *Natl. Wea. Dig.*, **28**, 12–24.
- , R. E. Jewell, and H. E. Brooks, 2002: Comparison between observed convective cloud-base heights and lifting condensation level for two different lifted parcels. *Wea. Forecasting*, **17**, 885–890.
- Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785.
- Davies, J. M., 2004: Estimations of CIN and LFC associated with tornadic and nontornadic supercells. *Wea. Forecasting*, **19**, 714–726.
- Donaldson, R. J., R. M. Dyer, and M. J. Krauss, 1975: An objective evaluator of techniques for predicting severe weather events. Preprints, *Ninth Conf. on Severe Local Storms*, Norman, OK, Amer. Meteor. Soc., 321–326.
- Donavon, R. A., and K. A. Jungbluth, 2007: Evaluation of a technique for radar identification of large hail across the Upper Midwest and central plains of the United States. *Wea. Forecasting*, **22**, 244–254.
- Doswell, C. A., III, and P. M. Markowski, 2004: Is buoyancy a relative quantity? *Mon. Wea. Rev.*, **132**, 853–863.
- , J. T. Schaefer, D. W. McCann, T. W. Schlatter, and H. B. Wobus, 1982: Thermodynamic analysis procedures at the National Severe Storms Forecast Center. Preprints, *Ninth Conf. on Weather Forecasting and Analysis*, Seattle, WA, Amer. Meteor. Soc., 304–309.
- , R. Davies Jones, and D. L. Keller, 1990: On summary measures of skill in rare-event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585.
- , H. E. Brooks, and M. P. Kay, 2005: Climatological estimates of daily local nontornadic severe thunderstorm probability for the United States. *Wea. Forecasting*, **20**, 577–595.
- Edwards, R., and R. L. Thompson, 1998: Nationwide comparisons of hail size with WSR-88D vertically integrated liquid water and derived thermodynamic sounding data. *Wea. Forecasting*, **13**, 277–285.
- Fawbush, E. F., and R. C. Miller, 1953: A method for forecasting hailstone size at the earth's surface. *Bull. Amer. Meteor. Soc.*, **34**, 235–244.
- Foster, D. S., and F. C. Bates, 1956: A hail size forecasting technique. *Bull. Amer. Meteor. Soc.*, **37**, 135–141.
- García-García, F., and R. List, 1992: Laboratory measurements and parameterizations of supercooled water skin temperatures and bulk properties of gyrating hailstones. *J. Atmos. Sci.*, **49**, 2058–2073.
- Guyer, J. L., and R. Ewald, 2004: Record hail event—Examination of the Aurora, Nebraska, supercell of 22 June 2003. Preprints, *22nd Conf. on Severe Local Storms*, Hyannis, MA, Amer.

- Meteor. Soc., 11B.1. [Available online at <http://ams.confex.com/ams/pdfpapers/82087.pdf>.]
- Hales, J. E., 1988: Improving the watch/warning system through use of significant event data. Preprints, *15th Conf. on Severe Local Storms*, Baltimore, MD, Amer. Meteor. Soc., 165–168.
- Hart, J. A., 1993: SVR PLOT: A new method of accessing and manipulating the NSSFC Severe Weather Database. Preprints, *17th Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 40–41.
- , J. Whistler, R. Lindsay, and M. Kay, 2003: NSHARP, version 3.90. Storm Prediction Center, National Centers for Environmental Prediction, Norman, OK.
- Kitzmilller, D. H., and J. P. Breidenbach, 1993: Probabilistic nowcasts of large hail based on volumetric reflectivity and storm environment characteristics. Preprints, *26th Int. Conf. on Radar Meteorology*, Norman, OK, Amer. Meteor. Soc., 157–159.
- Knight, C. A., and N. C. Knight, 2001: Hailstorms. *Severe Convective Storms, Meteor. Monogr.*, No. 50, Amer. Meteor. Soc., 223–254.
- Matson, R. J., and A. W. Huggins, 1980: The direct measurement of the sizes, shapes and kinematics of falling hailstones. *J. Atmos. Sci.*, **37**, 1107–1125.
- McCaul, E. W., and M. L. Weisman, 2001: The sensitivity of simulated supercell structure and intensity to variations in the shapes of environmental buoyancy and shear profiles. *Mon. Wea. Rev.*, **129**, 664–687.
- Miller, R. C., 1972: Notes on analysis and severe storm forecasting procedures of the Air Force Global Weather Center. AWS Tech. Rep. 200 (Rev.), Headquarters, Air Weather Service, Scott AFB, IL, 106 pp.
- Moore, J. T., and J. P. Pino, 1990: An interactive method for estimating maximum hailstone size from forecast soundings. *Wea. Forecasting*, **5**, 508–526.
- Morris, D. A., and P. R. Janish, 1996: The utility of mesoscale versus synoptic scale surface observations during the Lahoma hail and windstorm of 17 August 1994. Preprints, *18th Conf. on Severe Local Storms*, San Francisco, CA, Amer. Meteor. Soc., 60–64.
- NCDC, 2000: *Storm Data*. Vol. 42, 131 pp.
- NOAA, cited 2008: AWIPS D-2D user manual. [Available online at http://www-md.fsl.noaa.gov/~cheatwoo/AWIPS/OB4-UM/6_2_2.htm.]
- NSSL, cited 2008: Understanding damage and impacts. [Available online at http://www.nssl.noaa.gov/primer/hail/hail_damage.html.]
- Poolman, E. R., 1992: Die voorspelling van haelkorrelgroei in Suid-Afrika (The forecasting of hail growth in South Africa). M.S. thesis, Faculty of Engineering, University of Pretoria, 113 pp.
- Renick, J. H., and J. B. Maxwell, 1977: Forecasting hailfall in Alberta. *Hail: A Review of Hail Science and Hail Suppression, Meteor. Monogr.*, No. 38, Amer. Meteor. Soc., 145–151.
- Rotunno, R., and J. B. Klemp, 1985: On the rotation and propagation of simulated supercell thunderstorms. *J. Atmos. Sci.*, **42**, 271–292.
- Schaefer, J. T., J. J. Levit, S. J. Weiss, and D. W. McCarthy, 2004: The frequency of large hail over the contiguous United States. Preprints, *14th Conf. Applied Climatology*, Seattle, WA, Amer. Meteor. Soc., 3.3. [Available online at <http://ams.confex.com/ams/pdfpapers/69834.pdf>.]
- Smith, T. M., K. L. Ortega, and A. G. Kolodziej, 2007: Enhanced, high-density severe storm verification. Preprints, *23rd Conf. on Interactive Information Processing Systems*, San Antonio, TX, Amer. Meteor. Soc., 4B.3. [Available online at <http://ams.confex.com/ams/pdfpapers/120091.pdf>.]
- Thompson, R. L., C. M. Mead, and R. Edwards, 2007: Effective storm-relative helicity and bulk shear in supercell thunderstorm environments. *Wea. Forecasting*, **22**, 102–115.
- 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 shears. *Mon. Wea. Rev.*, **112**, 2479–2498.