DEVELOPMENT OF FORECAST EQUATIONS TO PREDICT THE SEVERITY OF HAIL EVENTS IN NEW YORK STATE

Kenneth D. LaPenta and George J. Maglaras

NOAA/National Weather Service
Albany, New York

Robert R. Mundschenk

NOAA/National Weather Service
Binghamton, New York

Abstract

Forecasters across the country routinely make subjective assessments of convective potential for their forecast area based on the values of various atmospheric parameters and indices. If convection is possible, forecasters must decide whether it will be severe or non severe; and if severe thunderstorms are possible, they must determine if the primary threat will be large hail, damaging winds, tornadoes, or all three. The specific parameter values which influence certain decisions may vary from person to person depending on a forecaster's geographic location, experience, and scientific understanding of the physical processes associated with thunderstorm development and evolution. Because of the subjective nature of the decision making process, the results may not be consistent. An equation developed in Maglaras and LaPenta (1997) provided guidance on forecasting tornadic, non-tornadic but severe and non-severe thunderstorm days. It didn't identify whether the main threat from non-tornadic severe storms was damaging winds or large hail (diameter .75 inch or larger). This paper is a follow-up study to Maglaras and LaPenta (1997) and describes the development of equations that provide objective statistical guidance for determining the overall severity of hail days by category, and the expected maximum hail size in New York State. The categorical forecast equation developed in this study successfully discriminates between major and minor hail days, and the hail-size equation showed skill at forecasting the maximum hail size for a day with thunderstorms. Although these equations should only be applied in the specific geographical area for which they were derived, the method used to develop these equations can be applied elsewhere.

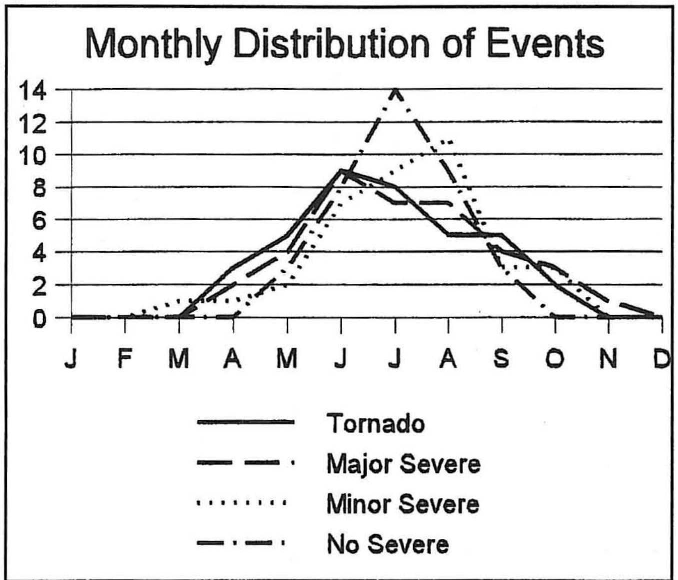
1. Introduction

One of the primary missions of National Weather Service (NWS) forecast offices is the issuance of tornado and severe thunderstorm warnings. Warnings are issued for occurring or imminent severe weather and thus have a short lead time. They are typically issued for a small geographical area (usually 1 or 2 counties) and for a duration of an hour or less. Tornado and severe thunderstorm watches for the entire country are issued by the NOAA/National Weather Service (NWS) Storm Prediction Center (SPC). Watches are issued for large geographical areas (parts of several states) and with lead times of several hours. These watches alert the public that general weather conditions are favorable for severe thunderstorms or tornadoes.

The identification of the synoptic and mesoscale meteorological conditions that are associated with tornadoes and severe thunderstorms typically is the initial step in the watch and warning process. Forecasters across the country routinely make subjective assessments of convective potential for their forecast area based on the values of various atmospheric parameters and indices used in conjunction with conceptual models. If convection is possible, forecasters must decide whether it will be severe or non-severe; and if severe thunderstorms are possible, they must determine if the primary threat will be large hail, damaging winds, tornadoes, or all three. The specific parameter values which influence certain decisions may vary from person to person depending on a forecaster's geographic location, experience, and scientific understanding of the physical processes associated with thunderstorm development and evolution.

Local NWS forecast offices have generally relied on the SPC for the assessment of hail potential and severity. SPC watch messages include a general hail size forecast (1 in., 2 in., 3 in., etc.). Pattern recognition, climatology and forecasted storm type are the primary input into SPC hail size forecasts (Johns and Doswell 1992). In this study, a method is developed to provide locally-based objective guidance for the assessment of hail potential and severity in advance of the issuance of severe weather watches. This locally-based guidance can also be valuable for assessing the need for additional staffing.

LaPenta and Maglaras (1993) (hereafter referred to as LM93) began a multi-step process to identify the general atmospheric conditions that were associated with days that featured severe thunderstorm events of various intensities in New York State. In the first step (LM93), an analysis was done to determine the general atmospheric conditions that were associated with tornadic thunderstorm days. In the second step (LaPenta 1995) (hereafter referred to as L95), an analysis was carried out to differ-



	Tor	Maj	Min	Non	Total
1989	4	8	4	2	18
1990	8	4	7	13	32
1991	4	11	12	10	37
1992	8	7	6	8	29
1993	13	7	8	4	32

Fig. 1. The monthly distribution of thunderstorm days used in this study for each of the 4 severe categories (upper portion). The lower portion shows the distribution of thunderstorm days by year for each of the 4 severe categories.

entiate the atmospheric conditions that produce tornadic thunderstorm days, major[1] severe thunderstorm days, and minor severe thunderstorm days. In the third step, Maglaras and LaPenta (1997) (hereafter referred to as ML97) described the development of an equation to make conditional forecasts of the severity of a thunderstorm day given the occurrence of thunderstorms. The equation developed in ML97 provides guidance on forecasting tornadic, non-tornadic but severe, and non-severe thunderstorm days. However, their equation didn't identify whether the main threat from non-tornadic severe storms would be damaging winds or large hail (diameter .75 inch or larger). This study uses the data set developed in the previous studies, LM93, L95 and ML97, to develop conditional forecast equations to predict the overall severity of hail days by category, and the expected maximum hail size in New York State. The purpose of these equations is to provide objective statistical guidance based on the forecaster's subjective assessment of the general atmospheric conditions expected at the time of the event.

2. Data

In LM93, the general atmospheric conditions that were associated with tornadoes in New York State on 24 days from 1989 to 1992 were examined. In L95, 111 days from 1989 to 1993 with severe weather in New York State were examined, 37 of which produced tornadoes. In that study, an analysis was carried out to differentiate the general atmospheric conditions that were associated with tornadic thunderstorm days, major severe thunderstorm days, and minor severe thunderstorm days. The data on the tornadic and severe thunderstorm days were obtained from *Storm Data* (U.S. Department of Commerce 1989-1993). In ML97 a statistical analysis was carried out to develop an equation to make conditional forecasts of the severity of a thunderstorm day given the occurrence of thunderstorms. The purpose of the equation was to provide objective statistical guidance to forecasters, using many of the methods and tools forecasters had been using for years to make subjective assessments of the potential for severe convection. The equation's objective output was based on the forecaster's subjective assessment of the general atmospheric conditions expected at the time of the event. That analysis was performed using thunderstorm data from L95 as part of the developmental sample. These data included 37 tornadic thunderstorm days, 37 major severe thunderstorm days, and 37 minor severe thunderstorm days. In order to include a sample of non-severe thunderstorm days, that data set (L95) was expanded to include an additional 37 days where thunderstorms occurred, but no severe weather was reported. The distributions of these thunderstorm days by month and by year are shown in Fig. 1. The equations developed in this study are based on the 148 days used in ML97 for the period from January 1989 through December 1993.

For each of the 148 days used, a sounding was constructed to approximate the synoptic-scale atmospheric conditions at the time of the event. Actual atmospheric soundings from across the northeastern United States were examined, and the sounding that was considered to be most representative of the airmass over the location where tornadoes, severe or non-severe thunderstorms occurred was selected. This sounding was then modified using the Skew-T Hodograph Analysis and Research Program (SHARP) (Hart and Korotky 1991) for observed surface temperature, dewpoint and wind from a surface observation site near the location and at the time of the thunderstorms. On a few occasions, additional subjective modifications were made if significant thermal advection aloft was evident, or changes to the vertical wind profile were warranted due to wind speed and/or direction changes aloft.

[1] Major severe weather days were defined as those days with 10 or more severe weather reports in the northeastern United States, and minor severe weather days were defined as those days with less than 10 events. The northeastern United States as defined in L95, ML97 and in this paper includes New York, New England, New Jersey and Pennsylvania.

Table 1. Definitions of hail size categories and of number of event categories.

Hail size category	Hail Size (in.)	Event category	Number of events
1	< .75	1	0
2	.75 < 1.00	2	1 to 2
3	1.00 < 1.75	3	3 to 5
4	≥ 1.75	4	6 to 14
		5	≥ 15

Table 2. Overall hail severity category.

Sum of hail size and number of event categories (CAT)	Hail severity category
CAT < 3.5	no severe hail
CAT ≥ 3.5 and < 5.5	minor severe hail
CAT ≥ 5.5 and < 7.5	major severe hail
CAT ≥ 7.5	extreme hail event

The limited spatial and temporal sampling by the NWS radiosonde network and the highly variable nature of the atmosphere make it difficult to create soundings that accurately represent the state of the atmosphere at the time of a particular event. If temporal and spatial restrictions are too strict, it will be difficult to come up with a statistically significant number of cases (Brooks et al. 1994). The goal of this study was to evaluate the general conditions that produce thunderstorms with non-severe or severe hail using information that is routinely available to forecasters. In order to maximize the size of the data set, strict temporal and spatial constraints were not placed on the use of observed soundings. Atmospheric conditions at the time of an event, or series of events, were approximated to the best degree allowed given data limitations. However, prior to the final selection of the 148 cases used in this study, a number of events were eliminated from consideration, because missing or incomplete data made analysis of the event impossible. Brooks et al. (1994) discuss in detail the use of, and limitations of, such an approach.

3. Methodology

The categorical hail severity of a particular thunderstorm day is defined to be a function of both the maximum observed hail size and the number of reports of severe hail. Hail size[2] is often reported by comparison to certain physical objects (coins, peas, mothballs, golf balls) or often rounded to simple values (e.g., 0.75 inch, 1 inch, 2 inches). As a result, maximum observed hail sizes were clustered around particular values and are not entirely randomly distributed. For the purpose of forecasting the categorical severity of a hail day, maximum hail sizes were divided into four categories (Table 1), with hail size categories ranging from one (non-severe hail) to four (hail greater than or equal to 1.75 inches). The maximum hail size categories were subjectively chosen in order to distribute the number of hail days as evenly as possible into each of the four categories. Categories for the number of hail events on a given day (Table 1) varied from one (no observed severe hail) to five (more than 15 reports of severe hail). These categories were also subjectively chosen in order to distribute the hail days as evenly as possible into each of the five categories. For the number of hail events categories, no attempt was made to adjust the spatial or time resolution of the individual hail reports in Storm Data. As a result, all hail events reported for a particular day could have been the result of only one or two intense thunderstorms producing many reports of large hail across only one or two counties, or many thunderstorms spread across a large geographic area yielding only one or two hail reports per county.

The sum of the hail size category and the number of events category was used to determine the overall categorical severity of a hail episode; this resulted in the four category definitions of hail severity listed in Table 2. Based on the values shown in Table 1 and Table 2, a non-severe hail day was defined as a day with no reports of hail .75 inch or larger. A minor hail day was a day with either five or less reports of severe hail less than 1.00 inch, or one or two reports of hail 1.00 inch to less than 1.75 inches. A major hail day was defined as a day with six or more reports of severe hail less than 1.00 inch, 3 to 14 reports of hail 1.00 inch to less than 1.75 inches, or one or two reports of hail 1.75 inches or greater. An extreme hail day was defined as a day with six or more reports of hail 1.75 inches or greater, or more than 15 reports of hail 1.00 inch or greater. These categories were subjectively determined based on the principle that an extreme hail day must have both a large number of events and very large hail reported. A major hail day must be the result of a large number of reports of relatively small hail, or a few reports of very large hail. Finally, a minor hail day must be the result of a few reports of relatively small hail.

The Statistical COrrelation and REgression program (SCORE) (Wooldridge and Burrus 1995) and SYSTAT Version 7.0.1 (1997) were used to perform analyses on the 148 cases to determine what meteorological parameters were best correlated with the maximum hail size and the hail severity categories. A total of 24 meteorological parameters were initially evaluated as possible predictors. However, based on correlations within the developmental sample, the predictor set was reduced to 10. These predictors were the lifted index (LI), equilibrium level (EQLV), 0-6 km mean wind (non-density weighted), convective available potential energy (CAPE), total-totals index (TT), sweat index, energy-helicity index (EHI) (Hart and Korotky 1991), 0-3 km storm relative helicity (SRH), wet-bulb zero categorical deviation (WBZCAT), and 850-mb temperature (850T). The values for CAPE, LI, and EQLV used in this paper were calculated by lifting the most unstable parcel in the lowest 150 mb (almost always the surface parcel). The WBZCAT was defined as shown in Table 3 and the reasons for using this approach

[2] Throughout this paper hail size refers to the diameter of hailstones.

Table 3. Wet-bulb zero categories.

Wet-bulb zero (ft)	Wet-bulb zero category
≤ 8000	2
8100 - 9000	1
9100 - 10900	0
11000 - 11900	1
12000 - 12900	2
13000 - 13900	3
≥ 14000	4

will be discussed later in the paper. The storm motion used to calculate the SRH was determined primarily from radar data. However, on the few occasions when radar data was not available, the storm motion was estimated using SHARP default storm motion and the text of NWS warnings and statements.

The 24 initial meteorological parameters that were examined were all parameters calculated by SHARP. They were chosen based on their applicability to the problem of hail forecasting. Only parameters included in the SHARP program were considered as candidates because of the requirement that all parameters used for equation development be readily available on a daily basis, both for operational use, and for the construction of the 148 soundings used for equation development. The other 14 meteorological parameters included the Bulk Richardson number, freezing level, deviation of the freezing level, deviation of the -20°C level, precipitable water, Theta E Index, 700-500 mb lapse rate, wet-bulb zero, storm-relative inflow, thickness of the 0°C to -20°C layer, height of the -20°C level, 500-mb temperature, 500-mb dewpoint depression and the 0-2 km storm relative moisture flux.

From the 148 cases, 25 were randomly selected to be used as a validation data sample. The remaining 123 days were used as the development data sample for the actual forward screening regression analysis. Based on the definitions of the severe weather categories in ML97, the validation sample included seven tornadic events, six major severe weather events, six minor severe weather events, and six days with no severe weather. Regression equations for the hail severity category and maximum hail size were determined and the validation data set was used to test these equations. Three case studies are also presented to illustrate the application of the predictor equations.

4. Regression Analysis Results

Table 4 shows the correlation of the 10 predictors used with the predictand data samples for the hail severity categories and the maximum hail size. The CAPE was the predictor most highly correlated with both predictands, followed closely by the LI. The EHI was the third most correlated predictor. The least correlated predictor was the WBZCAT.

Based on numerous iterations of the statistical software, it was determined that the best possible equations

Table 4. Correlation of the 10 predictors with the predictand data sample for the overall hail severity category and the maximum hail size.

	Hail Severity	Maximum Hail Size
Lifted Index (LI) (°C)	-.5340	-.4574
Equilibrium Level (EQLV) (k-ft)	.3175	.3388
Mean Wind (kt)	.2890	.3083
Convective Available Potential Energy (CAPE) (J kg^{-1})	.5418	.5119
Total Totals Index (TT) (°C)	.3093	.2923
Sweat Index	.3121	.3076
Energy-Helicity Index (EHI)	.4020	.4083
Storm-relative Helicity (SRH) (m^2 s^{-2})	.3069	.3741
Wet-bulb Zero Category (WBZCAT)	-.1693	-.1198
850-mb Temperature (850T) (°C)	.3237	.3178

for predicting the hail severity category and the maximum hail size (equations that maximized the number of correct forecasts on the validation data sample) were six-term equations, which included the CAPE, SRH, TT, 850T, EQLV (in thousands of ft), and the WBZCAT. The equation for overall hail severity (CAT) was:

$$CAT = -.144(EQLV) - .502(WBZCAT) + .00182(CAPE) + .0804(TT) + .00605(SRH) + .203(850T) + .153 \quad (1)$$

If CAT was less than 3.5, a non-severe hail day was forecast. If CAT was greater than or equal to 3.5 but less than 5.5, a minor severe hail day was forecast. For CAT greater than or equal to 5.5 but less than 7.5, a major severe hail day was predicted, and for CAT greater than or equal to 7.5, an extreme event day was forecast. The hail severity categories are summarized in Table 2.

The forecast equation for maximum hail size (SIZE) in inches was:

$$SIZE = -.0318(EQLV) + .000483(CAPE) + .0235(TT) + .00233\,(SRH) - .124\,(WBZCAT) + .0548\,(850T) - .772 \quad (2)$$

a. Physical basis of predictors

CAPE and LI were found to be the 2 variables best correlated with hail size and the hail severity categories. This is not surprising since the formation of severe hail is dependent on an updraft strong enough to support the weight of a hailstone long enough to allow growth to .75 inch diameter or larger, and updraft strength is strongly dependent on atmospheric instability (Johns and Doswell 1992). CAPE represents the total buoyancy of a rising air parcel integrated from the level of free convection to the equilibrium level. The LI estimates instability by comparing the temperature of a rising parcel to the temperature of its environment at only a single level (usually 500 mb). As a result, CAPE generally gives a better approximation of atmospheric instability. As the most highly correlated predictor, the CAPE was selected first by the regression program for both equations. However, once the CAPE was selected first, the LI became unimportant (despite the high correlation to the development sample) and does not appear in either equation because it substantially duplicates the information presented by CAPE (these two variables were also highly correlated with each other). Similarly, the EHI was highly correlated to the development samples, but it does not appear in the equation because the information it provides has already been accounted for by the CAPE and SRH. In contrast, the WBZCAT was the least correlated predictor individually, but was selected as the third predictor by the iterative regression process (after the CAPE and SRH had already been selected), because it had virtually no correlation with the other predictors and, thus, the information it provided had not already been accounted for.

Observations and numerical simulations indicate wind shear is important in organizing convection. (e.g., Fawbush and Miller 1954; Weisman and Klemp 1984). Given sufficient instability, increasing wind shear favors more organized multicellular and supercellular storms (Weisman and Klemp 1986). Observational experience suggests that very large hail is usually associated with supercells and well-organized multicellular systems (Johns and Doswell 1992). In addition, vertical shear may contribute to updraft strength. Brooks and Wilhelmson (1990) used numerical simulations to examine the relationship of updraft intensity to curvature shear of the environmental winds as measured by SRH. They concluded that this low-level shear as measured by SRH enhanced updraft intensity.

The thermal structure of the troposphere plays a role in the occurrence and size of hail at the surface of the earth. Melting of a hailstone as it falls is related to its initial size, the depth of the above-freezing layer through which it falls and the temperature of this layer. The WBZ level approximates the height of the freezing level in the downdraft air of the thunderstorm. The higher the WBZ level and the warmer the air below it, the greater the melting and the lower the likelihood that large hailstones would reach the ground (Johns and Doswell 1992). The lower the WBZ level and the colder the air below it, the less melting. However, WBZ levels that are too low usually are indicative of atmospheric conditions that are not favorable for severe convection. Even if severe convection does occur, the convection will be shallow and not be conducive to the formation of large hailstones. Therefore, Miller (1972) felt that the probability of hail was greatest when the WBZ was within a certain range. In this study the WBZ (and the actual environmental freezing level) showed little correlation to hail size or hail severity category. However, the deviation of the WBZ from a given value was examined and this significantly improved the correlation. Using the deviation from 10,000 ft provided the best results. (The average WBZ for the development data cases with severe hail was 10,900 ft). Forecaster experience indicated that an observed WBZ below this preferred level was not as detrimental as a deviation above the preferred level. This is because there would be less time for a falling hailstone to melt when WBZ values exhibit a negative deviation. WBZ categories based on the deviation from an ideal WBZ level were created (Table 3), and these values were used in the forecast equation.

b. Evaluation of the prediction equations

An evaluation of the equations on the development sample of 123 cases revealed the following. The CAT equation showed an overall correlation with the development sample of .50, while the correlation of the SIZE equation was .46. As a result, the equations account for 25% (or nearly 25% percent) of the variance of the dependent variable for the CAT (SIZE) equation. The equations were tested for significance using the F-test, and the F-test showed that the regression analysis results were significant at the 99% confidence level for both equations. The predicted mean value of the CAT (SIZE) equation was 5.04 (1.034), with a mean square error of 1.560 (0.509). The residuals showed independence with near zero mean (a 0.096 mean for the CAT equation and a -0.060 mean for the SIZE equation). In addition, the residuals were evenly distributed around the mean.

The equations were also evaluated on a validation sample of 25 cases. Using the categories provided in Table 2, the test results showed that the CAT equation was able to forecast the correct hail category 13 out of 25 times. Of the 12 incorrect forecasts, none were incorrect by more than one category.

The validation test results for the SIZE forecasts showed an average absolute error of .44 inches. Stratifying the validation sample into the definitions of the severe weather categories used in ML97, the average absolute hail size error for tornadic thunderstorm days

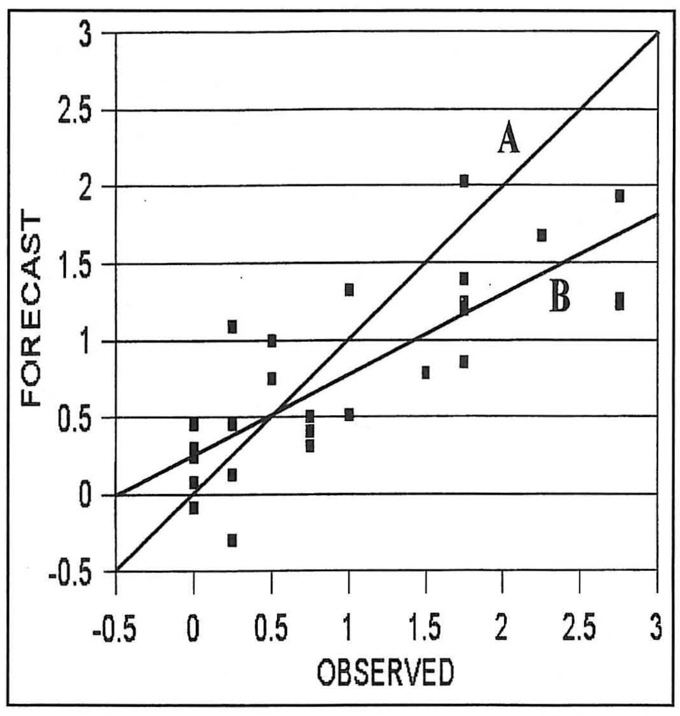


Fig. 2. Scatter diagram of forecast versus observed hail size for the validation data sample and case studies. Line A represents the line of best fit for a hypothetical forecast equation that has no bias. Line B is the actual line of best fit for the 28 cases plotted.

was .68 inches, for major severe weather days it was .43 inches, and for minor severe weather days it was .38 inches. Figure 2 is a scatter plot of the forecast hail size to the observed hail size for 28 cases (the 25 validation cases plus three additional case study days which will be discussed in section 6). Line A in Fig. 2 represents the line of best fit for a hypothetical maximum hail size equation that has no bias. Line B is the actual line of best fit for the 28 cases plotted in Fig. 2. Line B indicates that the SIZE equation generally underforecasts the maximum hail size when the observed hail diameter is 0.75 in. or greater, and that the underforecast bias increases with observed hail size. In addition, there is a tendency to overforecast hail size for non-severe hail.

An analysis of the categorical forecasts of hail severity for the 25 validation sample days and the three case study days was done in order to determine if the forecasts were better than random chance. The Heidke skill score for these 28 forecasts was 0.35 (random chance would produce a score of zero, while a perfect score would be one). Based on 28 forecasts and a four category forecast matrix, a Heidke skill score of zero suggests that between seven and eight forecasts would be correct by random chance. If the forecast equation was perfect and the Heidke skill score was one, then all 28 forecasts would be correct. In this case, the Heidke skill score was 0.35 with 16 of the 28 forecasts correct.

5. Operational Use of the Forecast Equations

For several years, forecasters have used the SHARP software application to modify actual atmospheric soundings in order to make subjective assessments of convective potential. First, based on their assessment of the general atmospheric conditions expected at a given time, they would determine the likelihood of thunderstorms forming. Second, they would determine the potential for any thunderstorms that did form to become severe. More recently, output from numerical model forecast soundings has become available to field forecasters. Using the SHARP application, forecasters can now analyze model-forecast soundings and make subjective assessments of the potential for convection, and they can do so as much as 48 hours in advance. The forecaster can accept the model sounding or make modifications to it for model biases, local effects or observed data. In addition, forecasters must also be aware of the effects of model generated convection on the predicted thermodynamic profiles. If the convective parameterization is triggered too quickly (too late) in the model, lapse rates and measures of stability will tend to be underforecast (overforecast) for a given forecast hour.

Operationally, it is generally more useful to use SHARP-derived data from model forecast soundings as input to the equations (Hart et al. 1998). Model soundings provide objective assessments of temperature, moisture and wind profiles valid for the exact time the forecaster is interested in. Observed atmospheric soundings can also be used to provide input into the equations. However, since observed atmospheric soundings are typically used to forecast potential severe weather 6 to 12 hours after the sounding observation time, they may require more extensive subjective modification than a model sounding.

The CAT and SIZE equations were developed using a sample of modified soundings, constructed to approximate the general synoptic scale atmospheric conditions at the time of the event. During operations, output from application programs (such as SHARP) that examine numerical model soundings or actual soundings is used as input to the equations in order to make objective conditional categorical forecasts of hail severity and maximum hail size. Because the equations are conditional in nature, they provide guidance on the overall hail severity and the maximum size of the hail, given that thunderstorms actually occur.

Operationally, forecasters can use these equations to determine the hail threat for a particular location, or, by applying these equations to model soundings at a number of locations, forecasters can attempt to identify the location within their forecast area most prone to severe hail. Forecasters can use these equations well in advance of the initiation of convection in order to provide the public with sufficient advance notice of possible severe weather, and to help determine the need for extra staffing. They can also use the equations in real time and make adjustments to model or actual soundings based on observed values such as surface temperature and dewpoint, wind structure based on WSR-88D wind profiles or surface data, and observed storm motion. When used in real time and with observed values, forecasters can use the equations to determine if their initial evaluation of the severe weather potential was correct, and, if not, make last minute adjustments to public forecast products and staffing needs.

6. Case Studies

As an additional test of the equations' ability to provide useful guidance to forecasters, three extreme hail days that were not part of the 25 day validation data set were examined. The first case, 29 May 1995, included widespread severe weather across the northeast portion of the U. S. including an F3 tornado. On that day, thunderstorms, including supercells, developed in a very unstable (CAPE 2489 J kg^{-1}) and highly sheared (SRH 324 m^2 s^{-2}) environment as shown in Fig. 3a. There were 67 reports (U.S. Department of Commerce 1995) of severe hail across the Northeast. The CAT equation correctly predicted an extreme hail day. While CAPE, TT, and EQLV had the greatest influence on the forecast of an extreme hail event, SRH and 850T also had a significant influence on the forecast. The SIZE equation forecast a maximum hail size of 2.02 inches as compared to the observed maximum hail size of 1.75 inches. Table 5 summarizes the values of the meteorological parameters input into the equations. Table 6 gives the value of each term in the CAT equation. Table 7 gives the value of each term in the SIZE equation.

The CAT equation again performed well for the 20 June 1995 case. On that day, the CAT equation correctly predicted an extreme hail day. There were 33 reports of large hail and 10 reports of damaging winds (U.S. Department of Commerce 1995). While atmospher-

Table 5. Predictors values for 3 case studies.

	Case 1 29 May 1995	Case 2 20 June 1995	Case 3 11 July 1995
EQLV (k ft)	40.7	46.2	36.5
WBZCAT	0	1	0
CAPE	2489	3915	2931
TT	55	53	50
SRH	324	-2	-10
850T	15.3	21.4	11.6

Table 6. Values of each term in the categorical hail severity forecast (CAT) equation, forecast and observed severity for the three case studies.

	Case 1 29 May 1995	Case 2 20 June 1995	Case 3 11 July 1995
EQLV (k ft)	-5.86	-6.65	-5.26
WBZCAT	0.00	-0.50	0.00
CAPE	+4.53	+7.13	+5.33
TT	+4.42	+4.26	+4.02
SRH	+1.96	-0.01	-0.06
850T	+3.11	+4.34	+2.35
Constant	+0.15	+0.15	+0.15
CAT	8.31	8.72	6.53
Forecast	extreme hail event	extreme hail event	major hail event
Observed	extreme hail event	extreme hail event	extreme hail event

Table 7. Values of each term in the forecast hail size equation, forecast maximum hail size (SIZE) and observed maximum hail size for the 3 case studies.

	Case 1 29 May 1995	Case 2 20 June 1995	Case 3 11 July 1995
EQLV (k ft)	-1.29	-1.47	-1.16
WBZCAT	0.00	-0.12	0.00
CAPE	+1.20	+1.89	+1.42
TT	+1.29	+1.25	+1.18
SRH	+0.75	0.00	-0.02
850T	+0.84	+1.17	+0.64
Constant	-0.77	-0.77	-0.77
SIZE	2.02 in.	1.95 in.	1.29 in.
Observed max hail size	1.75 in.	2.75 in.	2.75 in.

ic buoyancy was extreme (CAPE 3915 J kg^{-1}), environmental winds were not strong (40 kt or less through the troposphere) and low-level wind shear was small (SRH -2 m^2 s^{-2}) (Fig. 3b). CAPE and EQLV had the most influence on the forecast of an extreme hail day, with TT and 850T also making significant contributions to the forecast. The maximum size hail forecast was 1.95 inches with hail observed as large as 2.75 inches. Meteorological parameters, CAT equation terms, and SIZE equation terms are provided in Tables 5, 6 and 7 respectively.

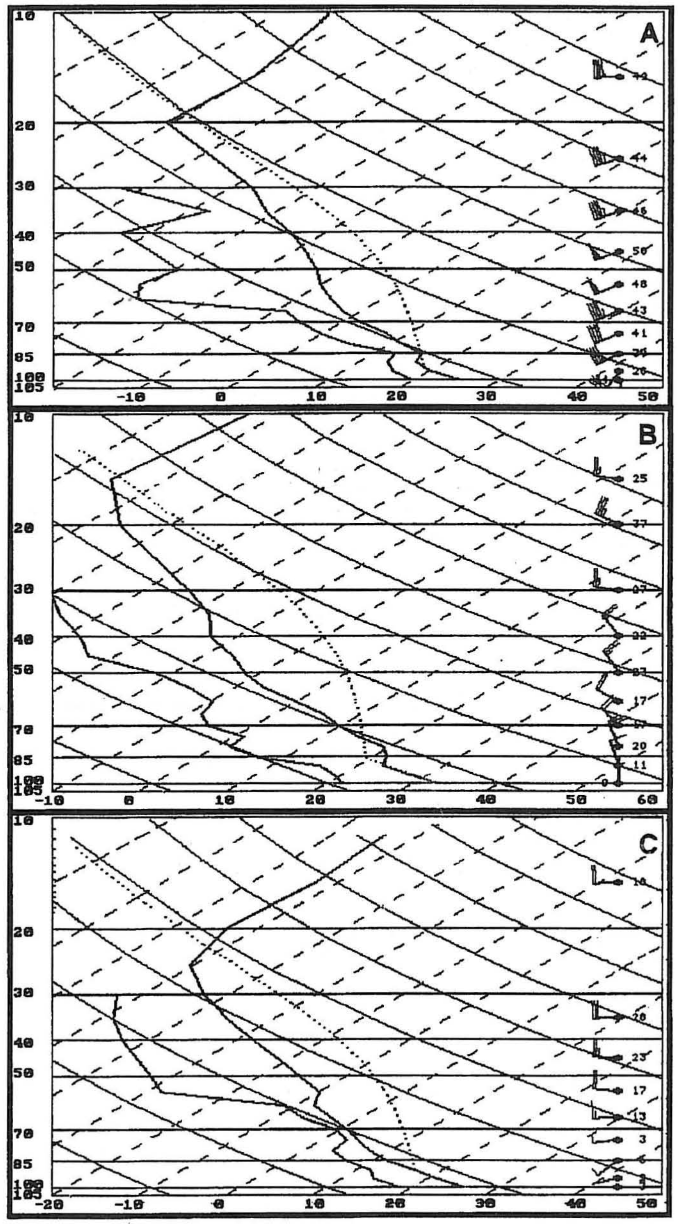


Fig. 3. Soundings for 29 May 1995 (A); 20 June 1995 (B); and 11 July 1995 (C) with the horizontal axis temperature (°C) and vertical axis pressure (mb x 10). Solid lines represent environmental temperature (right) and dewpoint (left). The dotted line indicates the path of a surface-based lifted parcel. Wind barbs (kt) along the right side are plotted using the standard convention.

The 11 July 1995 case was similar to the 20 June 1995 case in that hail was the predominant form of severe weather. There were 22 reports of large hail in the northeastern U. S. and just 4 reports of wind damage (U.S. Department of Commerce 1995). The atmosphere was quite unstable (CAPE 2931 J kg^{-1}), but not as unstable as the 20 June 1995 case. Winds were relatively light through the troposphere with minimal low-level wind shear (SRH = -10 $m^2 s^{-2}$) (Fig. 3c). The forecast equations both underestimated the severity of this hail day. The CAT equation forecast a major hail day as compared to the observed extreme day. The SIZE equation forecast of 1.29 inches was well below the maximum observed size of 2.75 inches. Several factors may have contributed to the error in the forecast. First, as discussed, the equation for SIZE is biased toward underforecasting hail size (Fig. 2). In addition, inaccuracies in the sounding created to describe atmospheric conditions on 11 July may have contributed to errors in the forecast equations. Third, standard sounding data provide only limited information that is relevant to hail formation processes, especially microphysical aspects of hail formation such as drop size distributions and entrainment rates. Finally, since hail size is usually estimated rather than measured directly, there is uncertainty in the accuracy of the hail size data base, and any inaccuracies can lead to equation forecast errors. Meteorological parameters, CAT equation terms, and SIZE equation terms are provided in Tables 5, 6 and 7 respectively.

7. Discussion

Forecasters at the NWS forecast office at Albany have found that the equation developed in ML97, which provides guidance on forecasting days when tornadic, non-tornadic but severe, and non-severe thunderstorms occur, to be a useful tool in assessing the potential for severe convection. However, the equation didn't identify whether the main threat from non-tornadic severe storms was damaging winds or large hail. This update to that study developed equations to forecast the categorical severity of a hail day and the maximum hail size. The CAT equation showed an overall correlation with the development data of .50 while the overall correlation of the SIZE equation was .46. These regression analysis results were tested for significance using the F-test, and the F-test showed that the regression analysis results were significant at the 99% confidence level for both equations. Based on the validation data, CAT equation forecasts of hail severity showed a Heidke skill score of 0.35. In comparison, the convective potential equation developed in ML97 showed greater skill with an overall correlation to the development data sample of .64, and validation data forecast results producing a Heidke skill score of 0.48. Despite the lower scores for the hail equations, the validation test results for maximum hail size and their graphical representation in Fig. 2 are similar to the forecast of maximum hail size results achieved by Moore and Pino (1990), and much better than the results achieved by Fawbush and Miller (1953).

The conditional nature of the two hail equations necessitates a two step approach to its application. First, forecasters must assess the likelihood that deep convection will develop. The output from these equations is not intended to provide any guidance for this forecast problem. Thus, if thunderstorms are not expected or do not form, then the equations' output has no meaning. If analyses of observed data and numerical model output indicate thunderstorms are possible, or if thunderstorms are already occurring, then the equations' output should provide useful guidance. In addition, when using the equations with numerical model output, systematic errors of the model will be reflected as systematic errors in the forecasts from the equations. However, if forecasters are aware of model errors or biases for their area, they can subjectively adjust the model output, thereby reducing the impact of this limitation. Finally, even though these equations should only be applied in the specific geographic area for which they were derived, the methods used to develop them can be applied elsewhere.

Acknowledgments

The authors would like to thank Laurie Hogan of the NWS Eastern Region Headquarters for her constructive review of this paper. Thanks also to David Kitzmiller of the NWS Techniques Development Laboratory, and Steven Weiss of the NWS Storm Prediction Center for their thorough reviews of this paper. Data for this study came from the NWS Albany Meso-Climatological Project's data archives.

Authors

Kenneth LaPenta has been a Lead Forecaster at the NWS Forecast Office at Albany, New York since 1986. Prior to joining the NWS in 1975 he worked at the National Hurricane Research Laboratory in Coral Cables, Florida and with Stone and Webster Engineering Corporation in Boston, Massachusetts. He earned a B.S. in Meteorology in 1971, and a M.S. in Meteorology in 1973 from Saint Louis University. His primary research interests are the analysis and forecasting of severe storms in the northeast U.S. Mr. LaPenta shared the NWA Operational Achievement Group Award for the year 2000 with Jonathan Blaes for outstanding teamwork in issuing early special weather statements and warnings resulting in no lives lost from the tornadoes that hit New York's Montgomery and Saratoga Counties on 31 May 1998.

George Maglaras has been a Lead Forecaster at the NWS Forecast Office at Albany, New York since 1988. He is interested in statistical forecast methods and analysis, and in forecast verification. From 1980 to 1986 he worked at the Techniques Development Laboratory in Silver Spring, Maryland. Between 1986 and 1988 he worked at the NWS Forecast Office in Washington, D.C., where he was part of the Satellite Field Service Station. He earned his B.S. Degree in Meteorology in 1978, and a M.S. Degree in Computer Science in 1981 from the City College of the City University of New York. In addition, between 1984 and 1986 he completed all course work toward a M.S. Degree in Meteorology at the University of Maryland.

Robert Mundschenk earned his B.S. degree in Meteorology in 1985 from the State University of New York at Oswego. He began his National Weather Service career in 1989 at the NWS Forecast Office in Buffalo, New York. He was a forecaster at the NWS Forecast Office at Albany, New York, from 1995 to 1998, and is currently a forecaster at the NWS Office at Binghamton, New York. He is interested in radar meteorology and lake-effect snow forecasting.

References

Brooks, H. E., and R. B. Wilhelmson, 1990: The effect of low-level hodograph curvature on supercell structure. Preprints, *16th Conference on Severe Local Storms*, Kananaskis Park, Alberta, Amer. Meteor. Soc., 588-592.

________, C. A. Doswell III and J. Cooper, 1994: On the environments of tornadic and non-tornadic mesocyclones. *Wea. Forecasting*, 9, 606-618.

Fawbush, E. J., and R. C. Miller, 1953: A method for forecasting hailstone size at the Earth's surface. *Bull. Amer. Meteor. Soc.*, 34, 235-244.

________, and ________, 1954: The types of airmasses in which North American tornadoes form. *Bull. Amer. Meteor. Soc.*, 35, 413-417.

Hart, J. A., and W. Korotky, 1991: The SHARP workstation v1.50 users guide. National Weather Service, NOAA, U.S. Department of Commerce, 30 pp. [Available from NWS Eastern Region Headquarters, Scientific Services Division, 630 Johnson Ave., Bohemia, NY 11716.]

Hart, R. E., G. S. Forbes, and R. H. Grumm, 1998: The use of hourly model-generated soundings to forecast mesoscale phenomena. Part I: Initial assessment in forecasting warm-season phenomena. *Wea. Forecasting*, 13, 1165-1185.

Johns, R. H., and C. A. Doswell III, 1992: Severe local storms forecasting. *Wea. Forecasting*, 7, 588-612.

LaPenta, K. D., 1995: Forecasting tornadic versus non-tornadic severe thunderstorms in New York State. *Eastern Region Technical Attachment, No. 95-4A*, National Weather Service, NOAA, U.S. Department of Commerce, 15 pp.

________, and G. J. Maglaras, 1993: New York State Tornadoes. *Eastern Region Technical Attachment, No. 93-12A*, National Weather Service, NOAA, U.S. Department of Commerce, 13 pp.

Maglaras, G. J., and K. D. LaPenta, 1997: Development of a forecast equation to predict the severity of thunderstorm events in New York State. *Natl. Wea. Dig.*, 21:3 , 3-9.

Miller, R. C., 1972: Notes on analysis and severe storm forecasting procedures of the Air Force Global Weather Central. TR 200 (Rev.), Air Weather Service (MAC), U.S. Air Force.

Moore, J. T. and J. P. Pino, 1990: An interactive method for estimating maximum hailstone size from forecast soundings. *Wea. Forecasting*, 5, 508-525.

SYSTAT Version 7.0.1 for Windows, 1997: SPSS Inc, Chicago, IL.

U.S. Department of Commerce, 1989-1993, 1995: *Storm Data,* 31-37.

Weisman, M. L., and J. B. Klemp, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, 112, 2479-2498.

________, and ________, 1986: Characteristics of isolated convection. *Mesoscale Meteorology and Forecasting*, P. S. Ray, Ed., Amer. Meteor. Soc., 331-358.

Wooldridge, M., and S. Burrus, 1995: *Statistical COrrelation and Regression (SCORE)*. National Weather Service, NOAA, U.S. Department of Commerce.