VIL Density as a Hail Indicator

STEVEN A. AMBURN AND PETER L. WOLF

National Weather Service Office, Tulsa, Oklahoma

(Manuscript received 1 January 1996, in final form 5 September 1996)

ABSTRACT

In current severe thunderstorm warning operations, forecasters frequently use the vertically integrated liquid water content (VIL) product from the WSR-88D to estimate thunderstorm severity and, particularly, hail size. Since VIL varies greatly based on airmass characteristics, forecasters have typically determined a threshold VIL to be used for each new thunderstorm event. A product that is independent of airmass characteristics, and thus independent of season and geographic location, would be more desirable in an operational warning environment.

It has been observed that high-topped thunderstorms with high VILs do not always produce large hail. It has also been observed that low-topped thunderstorms with low VILs occasionally do produce large hail. However, the maximum reflectivity in both high-topped and low-topped thunderstorms is similar when both produce similar-sized hail. From this, it was hypothesized that dividing the VIL by the echo top would ‘‘normalize’’ the VIL and produce a common value, or range of values, for thunderstorms producing large hail, independent of airmass characteristics. This quotient is defined as VIL density in this study.

To test the hypothesis, thunderstorm VIL and echo tops were recorded over a wide range of airmass characteristics, and VIL density was calculated. The data were correlated to surface-based reports of hail. The results showed a substantial increase in severe hail (≥19 mm, ¾ in.) reports as VIL density increased above 3.5 g m^{-3}. At values greater than 4.0 g m^{-3}, virtually every thunderstorm produced severe-criteria hail, regardless of the actual VIL or the thunderstorm height. At values below 3.5 g m^{-3}, very few thunderstorms produced severe-criteria hail.

1. Introduction

Greene and Clark (1972) suggested that vertically integrated liquid water content (VIL) could be a useful tool for assessing the severe weather potential of thunderstorms. Unfortunately, the minimum VIL value that correlates to ground reports of severe hail varies greatly because of a substantial dependence on airmass characteristics, such as the vertical profile of temperature and moisture. Identifying the appropriate VIL for a given day depends on several assumptions, which can lead to missed warning events and false alarms.

In an attempt to improve the warning program for severe hail and make greater use of the VIL product, a 9-month thunderstorm study was completed at the National Weather Service Office (NWSO) in Tulsa, using the KINX WSR-88D radar data at Inola, Oklahoma, about 20 mi east of Tulsa. Maximum VIL values and associated echo tops were used to calculate VIL densities (VIL/echo top). These VIL densities were correlated to surface reports below the thunderstorms to verify the presence of hail aloft. A total of 221 thunderstorms, from November 1994 through July 1995, were examined in the study. The majority of these produced severe hail.

The results indicated VIL density to be a useful hail indicator, without the airmass dependency of VIL alone and without the problematic assumptions of ‘‘VIL of the Day.’’ VIL density was also more effective than VIL of the day, especially for thunderstorms near the radar (within the ‘‘cone of silence’’). A threshold value of VIL density (3.5 g m^{-3}) was determined, which correctly identified 90% of the severe hail cases, even with widely varying VIL values, echo tops, and airmass characteristics. The design, associated problems, and results of this study are presented here. The problems associated with hail melting before it reaches the ground were not addressed in this study.

2. Vertically integrated liquid and reflectivity

VIL is a nonlinear function of reflectivity and converts weather radar reflectivity data into estimates of equivalent liquid water content based on theoretical studies of drop size distributions and empirical studies of reflectivity factor and liquid water content (U.S. Department of Commerce 1991). It should be emphasized that VIL is a radar-derived estimate of liquid water (ex-

---

Corresponding author address: Mr. Steven A. Amburn, National Weather Service, 10159 East 11th Street, Suite 300, Tulsa, OK 74128-3050.
E-mail: saa@nwstsa1.abrfc.noaa.gov

clusive of ice), based on numerous assumptions about reflectivity. However, it is important to add that VIL also varies according to hail size distribution and the refractive index at each elevation slice (height), since these factors also affect reflectivity. [For a complete discussion of these assumptions, the reader should refer to appropriate texts such as Rinehart (1991).] One form of VIL is given by the equation

$$\mathrm{VIL} = \sum 3.44 \times 10^{-6}[(Z_i + Z_{i+1})/2]^{4/7}\Delta h \quad (1)$$

and has units of kilograms per square meter (kg m^{-2}); Z_i and Z_{i+1} are radar reflectivity values ($\mathrm{mm}^6\ \mathrm{m}^{-3}$) at the lower and upper portions of the sampled layer; Δh is the vertical thickness of that layer in meters, which varies as a function of range and elevation scan strategies of the radar. For the WSR-88D, VIL is based on reflectivity over a discrete vertical thickness, which is summed for each 4 km × 4 km (2.2 nm × 2.2 nm) grid box (U.S. Department of Commerce 1991).

The reflectivity factor (Z) is proportional to the target diameter (D) to the sixth power and the total number of targets (n) within the measured volume and is given by

$$Z = \sum n_i \times D_i^6 \quad (2)$$

(Rinehart 1991). Equation (2) is only accurate for Rayleigh scattering, that is, target sizes up to about 32 mm (about 1.2 in.). Larger targets will backscatter in the Mie region and will in fact have a lower reflectivity than some smaller targets.

As indicated by the exponent, in Eq. (2), target diameter has a much greater effect on reflectivity than does the number of targets. In fact, the reflectivity for a single 19-mm target (¾ in.) is approximately 1000 times greater than for a single 6-mm target (¼ in.). This range of target sizes is significant in that raindrops in thunderstorms rarely exceed 6 mm (Houghton 1985), whereas hail can maintain sizes larger than 19 mm for tens of minutes. Because VIL increases with increasing reflectivity [based on Eqs. (1) and (2)], higher VIL values require higher reflectivity values, implying the presence of large targets, that is, large hail aloft. Figure 1 shows a plot of reflectivity factors for single targets of different sizes, with a significant increase in reflectivity occurring between 6 mm (a typical raindrop) and 19 mm (¾-in. hailstone).

It should be noted that the reflectivity factor has a large dependence on the hydrometeor phase, as well as target size. For the same-size spherical hydrometeor in the Rayleigh range, ice will have a lower reflectivity factor than water (Rinehart 1991). However, the WSR-88D algorithms assume all targets have the same dielectric constant (that for liquid water), so a 1-in. hailstone is treated as a 1-in. raindrop, resulting in a high reflectivity factor. Data throughout this study showed that thunderstorms that produced hail had higher reflectivity (and higher VIL) than those that did not produce hail.

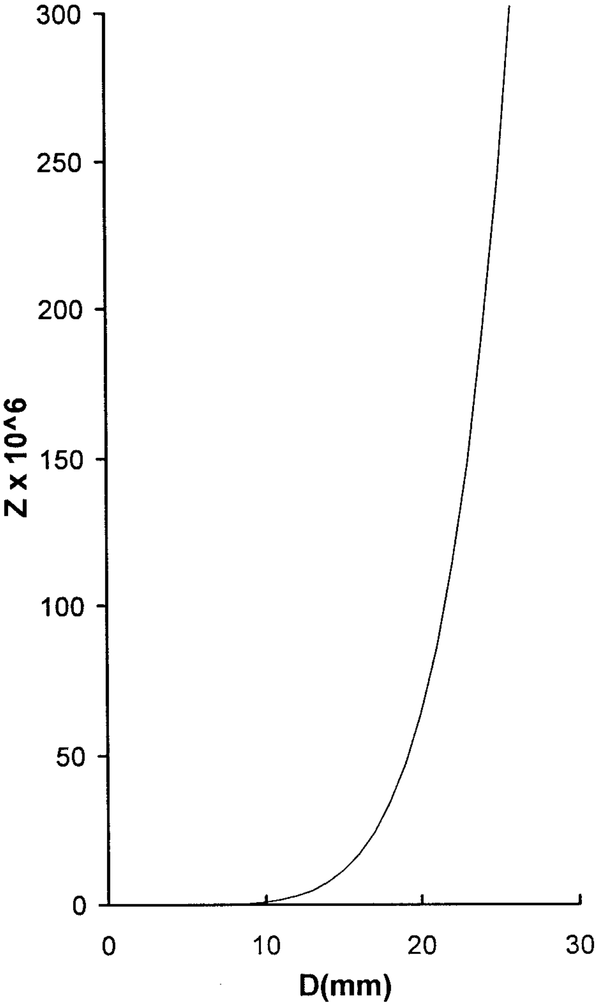

FIG. 1. Reflectivity (Z) versus target size (D) for a single target. Target diameter is shown in millimeters. Reflectivity is shown in millimeters to the sixth power. The curve shows a significant increase in reflectivity from a 6-mm raindrop to a minimally severe sized hailstone of 19 mm.

3. VIL of the day

Due to the correlation of VIL and hail, VIL is used to help identify thunderstorms that are likely to produce severe hail. However, warning forecasters discovered that thunderstorms in colder air masses could produce severe hail with low VILs (25–35 g m^{-2}), while those in warmer air masses often failed to produce severe hail with high VILs (50–60 g m^{-2}). To reduce the number of unwarned events, an appropriate VIL threshold for a particular day (VIL of the day) needed to be determined. A number of techniques have been tried with some success, including the use of temperatures aloft to determine the VIL of the day. Other National Weather Service offices have used different combinations of meteorological parameters. However, VIL of the day is based on assumptions that can result in problems.

First is the assumption that all thunderstorms within the radar umbrella will have the same characteristics. Given that assumption, each thunderstorm could be expected to have the same growth rate, the same maximum echo top, the same movement, and the same duration. Clearly, this is not the case. In fact, not every thunderstorm in the same air mass produces hail, much less the same size hail.

Second, VIL values must be reasonably accurate for the entire depth of the thunderstorm, if a VIL of the day is to be applied across the radar umbrella. This can be a significant problem for thunderstorms very close to the radar, that is, within the cone of silence, where the radar cannot sample the upper portion of the thunderstorm. In these instances, the VIL calculation is truncated at some altitude below the thunderstorm top, resulting in underestimated VIL values. These underestimates will rarely equal or exceed the VIL of the day, even though large hail may be present.

Finally, VIL of the day changes not only from one season to another, but often from one day to the next. Since VIL is a function of reflectivity throughout the height/depth of a particular thunderstorm, this generally results in low VILs for short storms (cold season) and high VILs for tall storms (warm season). Used in this manner, a tall thunderstorm with low overall reflectivity (small precipitation targets) may have the same VIL as that from a short thunderstorm with high reflectivity (large precipitation targets, i.e., severe hail).

An example of how VIL alone can be misleading was demonstrated on 17 May 1995. Two thunderstorms were located approximately equidistant from the radar and about 70 km (40 mi) apart. Hail 25.4 mm in diameter (1.0 in.) was observed from the northern thunderstorm, which had a VIL of 72 kg m^{-2} and an echo top of 15.24 km (50 000 ft). An appropriate VIL of the day might have been around 50 kg m^{-2}. However, at about the same time, 22.4-mm-diameter hail (0.88 in.) was reported from the southern thunderstorm with a VIL of only 37 kg m^{-2} and an echo top of 9.74 km (32 000 ft). Similar examples were common during the study.

4. VIL density

VIL density is simply the VIL (kg m^{-2}) divided by the echo top (m). The quotient is multiplied by 1000 to yield units of g m^{-3}:

$$\text{VIL density} = \text{VIL/echo top}. \tag{3}$$

Paxton and Shepherd (1993) described this ratio as ‘‘cell density’’ in a study using WSR-57 RADAP II data to locate all types of severe weather (hail, wind gusts, tornadoes) in Florida. For the NWSO Tulsa study, VIL density was correlated to hail alone because of the distinct mathematical relationship between target size and reflectivity [Eq. (2)].

When the VIL is ‘‘normalized’’ using the echo top, the resulting VIL density can be used to quickly identify thunderstorms with high reflectivities relative to their height. Those thunderstorms will often contain hail cores. As VIL density increases, the hail cores tend to be deeper and more intense, and reported hail sizes tend to be larger. The potential usefulness of VIL density for identifying thunderstorms containing large hail led to the NWSO Tulsa study.

5. Data

A total of 221 thunderstorms that occurred within range of the KINX WSR-88D from November 1994 through July 1995 were examined. Severe hail was produced by 185 of the thunderstorms, while the remaining 36 thunderstorms produced small hail or no hail at all. Hail diameters ranged from zero (‘‘no-hail’’ thunderstorms) to over 50 mm (2 in.). Thunderstorms ranged from 18.5 to 213 km (10–115 nm) from the radar; VIL values ranged from 17 to 91; radar-defined echo tops ranged from 6.7 to 18 km (22–59 000 ft).

Two criteria were used in selecting cases. First, all thunderstorms that were known to have produced severe hail, as indicated from local storm report logs, were included in the study. Second, thunderstorms that did not produce severe hail were included only if they moved over a highly populated area when reports of hail observations could normally be expected (from around 0800 until around 2200 LT). This was done to ensure that the thunderstorms included as nonsevere hail producers did not, in fact, produce severe-criteria hail. Thunderstorms that were not observed to produce severe hail, and that did not move over a highly populated area, were not included in the study. Also, thunderstorms with maximum VILs below 15 were not included in the study.

This method of data screening was used to solve two data collection problems. First, some high VIL (and high VIL density) thunderstorms may produce severe hail that goes unreported due to low population density (the ‘‘thunderstorm in the middle of nowhere’’). The lack of a surface reports of large hail does not necessarily mean large hail was not present aloft. Second, due to staffing and time constraints, it is typical and practical to not search for ‘‘nonreports.’’ Finally, thunderstorms with VILs of 15 or greater that did not produce hail were only included if they moved over a highly populated area. As a result, the database consists of only a small number of thunderstorms that did not produce severe hail.

For each thunderstorm producing severe-criteria hail, the maximum VIL during either of the two volume scans prior to the surface report was recorded. The WSR-88D echo top was also recorded for either the same pixel of that same volume scan, or the next pixel downstream, if greater, to account for problems due to storm motion or tilt. VIL Density was then calculated. For each thunderstorm that produced small hail or no hail at all, the maximum VIL density was calculated while the thunderstorm was over a populated area, using the highest value of VIL and lowest corresponding echo top for the same pixel.

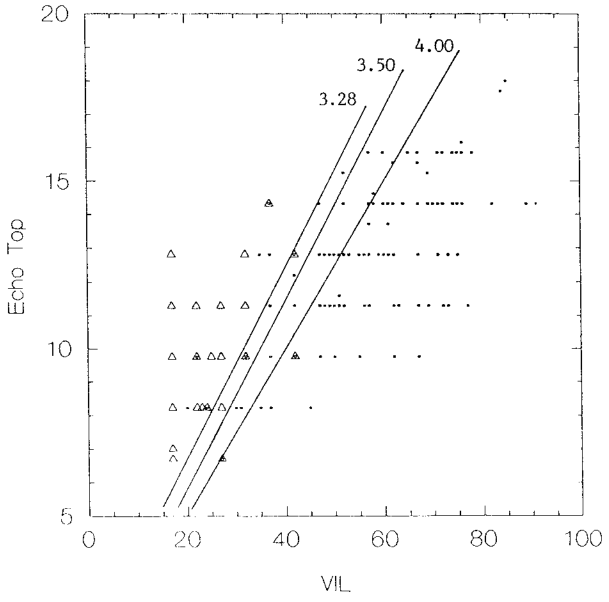

FIG. 2. Scatter diagram of echo top versus VIL for 185 severe hail cases (■) and 36 nonsevere hail cases (△). (Due to the number of data points, many are overplotted.) Values of VIL density (g m^{-3}) are shown as solid lines, labeled 3.28, 3.50, and 4.00. Echo top is shown in kilometers.

6. Results

The 221 cases in the study were stratified into groups of severe and nonsevere cases, and sorted for hail size, VIL, echo top, and VIL density. A VIL density threshold was identified, as suggested by the graph in Fig. 1, and statistics based on different threshold values were determined.

Figure 2 is a scatter diagram of echo top versus VIL for 185 severe hail cases and 36 nonsevere hail cases. Values of VIL density (g m^{-3}) are shown as solid lines, labeled 3.28, 3.50, and 4.00. The VIL density line labeled 3.28 initially appeared to be a reasonable threshold value, identifying 96.7% (179 of 185) of the severe hail cases. (The value 3.28 is obtained when the thunderstorm echo top in thousands of feet equals the VIL value in kg m^{-2}.) However, using 3.28 as a threshold would have falsely identified as severe 25% of the nonsevere hail cases (9 of 36). A closer examination of the results revealed that 90% of the thunderstorms that produced severe hail had a VIL density of 3.5 or greater. Conversely, the threshold of 3.5 falsely identified as severe less than 2% of the cases. When considering the probability of detection and false alarm rates, this makes 3.5 a more reasonable threshold value (see Fig. 3).

In addition, the study indicated that as VIL density increased, the maximum reported hail sizes also increased. This corresponds well with the common knowledge that thunderstorms with higher VILs often produce larger hailstones. The same relationship is true for VIL density and will be discussed more completely in section 8.

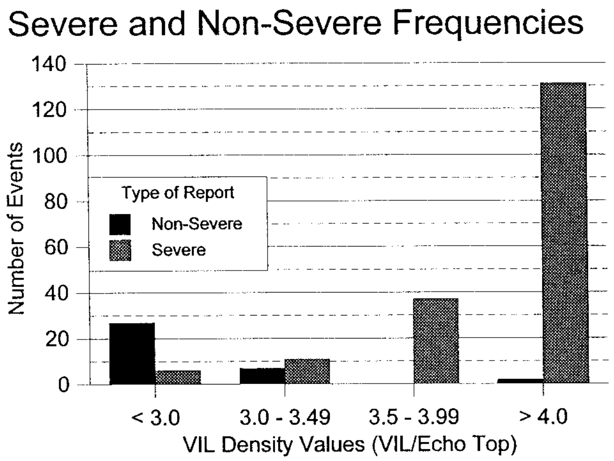

FIG. 3. The number of nonsevere and severe events are graphed as labeled, indicating a significant increase in the number of severe events as VIL density values increased above 3.5 g m^{-3}. For VIL density values of 3.5 g m^{-3} and greater, 168 thunderstorms produced severe-criteria hail, while only 2 did not.

The example at the end of section 3 described two thunderstorms with significantly different VILs where both thunderstorms produced severe-criteria hail. The VIL density for the northern thunderstorm was 4.72, while the southern thunderstorm had a VIL density of 3.79. Using the VIL density threshold value of 3.5 g m^{-3}, warnings would have been issued for both thunderstorms, with larger hail expected from the northern thunderstorm. Using VIL of the day would have resulted in a missed event for the southern thunderstorm.

7. Complexities

Several problems were encountered during the study that had an influence on the outcome. First, the technique for computing VIL can affect its value, depending on thunderstorm speed of movement and range. Second, the echo-top measurement is frequently not accurate due to the discrete elevation scan strategy. Third, verification practices are designed to efficiently verify warnings, not to satisfy scientific studies.

VIL is computed by integrating reflectivity for each vertical column of pixels. For slow-moving, vertical storms VIL calculations should be quite accurate. However, for fast-moving storms (or slow-moving, strongly tilted storms) VIL calculations will frequently not be accurate (U.S. Department of Commerce 1991). This error in calculation results when the hail shaft begins the volume scan in one pixel but moves into the downstream pixel before the volume scan is complete. For a hail shaft translating at the thunderstorm speed of 48 km h^{-1} (30 mph), the hail shaft will move 2.4 km (1.5 mi) in 3 min. If that hail shaft began in the middle of a 4 km × 4 km pixel, it would be 0.4 km (0.25 mi) into the next downstream pixel when the volume scan [VCP (volume coverage pattern) 21] is only half complete. Of course, results will be better or worse depending on both the thunderstorm speed and where the hail shaft was located in the pixel at the beginning of the volume scan. It is interesting to note that most of the low VIL density thunderstorms in this study that produced severe hail occurred in fast-moving bow-echo events, where storm motion and tilt were both maximized, resulting in underestimated VIL values.

For thunderstorms at some distant range, the VIL computation problem is diminished since the thunderstorm and hail shaft are normally sampled completely in the first few minutes of the volume scan, by the first few elevation slices and, therefore, within the same pixel. However, for very distant thunderstorms, VIL may be overestimated since the equivalent reflectivity value of the lowest elevation slice is assumed to extend to the ground.

It has been documented that WSR-88D echo tops can vary significantly with changes in range, even though the thunderstorm itself may not change in height (U.S. Department of Commerce 1991, part C). Only a finite number of elevation angles are provided by the WSR-88D radar. At NWSO Tulsa, VCP 21 is used to provide the best possible velocity measurements, despite its fewer number of elevation angles. However, because of "gaps" in the scan strategy of VCP 21, the radar does a poorer job measuring echo tops and estimating VIL than does VCP 11 (U.S. Department of Commerce 1991). Due to the emphasis on tornado detection at NWSO Tulsa, VCP 21 was used during the entire study.

Examples of echo-top estimates truncating between 30 and 50 dB*Z* were frequently observed in the reflectivity imagery products, indicating that echo tops were often underestimated. [The WSR-88D echo-top product uses the highest (in altitude) sample volume that meets the minimum reflectivity value of 18.5 dBZ.] This truncation problem can be quite extreme near the radar (within the cone of silence), where the radar may sample only the lowest portions of the thunderstorm. However, the calculation of VIL density is not affected by the truncation problem, since VIL density normalizes the reflectivity of thunderstorms, regardless of their overall height.

Another complicating factor concerned the warning verification practices. Verification telephone calls to locate weather events often stop after the first severe weather report is received. Additional calls to determine the largest severe hail observed are not normally made.

8. VIL density and hail size

An interesting outcome of the study was that forecasters began to subjectively use VIL density to estimate hail sizes for use in their warnings. Stratification of the data did reveal an apparent correlation, indicating it might be possible to estimate "expected" hail sizes based on VIL density. Table 1 shows the distribution of reported hail sizes for given ranges of VIL density.

TABLE 1. Hail sizes for given ranges of VIL density. Note that as the hail sizes increase, the minimum ranges for VIL density also increase. Values in parentheses represent the number of events in that category–range.

	Ranges in VIL density					
Hail size	<3.0	3.0–3.4	3.5–3.9	4.0–4.4	4.5–4.9	>4.9
<19 mm (36)	27	7	0	2	0	0
19–24 mm (117)	6	10	32	44	18	7
25–45 mm (63)	0	1	5	18	16	23
>45 mm (5)	0	0	0	0	1	4
(Total number)	(33)	(18)	(37)	(64)	(35)	(34)

Due to the method of verification, explained earlier, the largest hail to reach the ground is often not found. However, the data did show that for the range of VIL density from 3.5 through 3.9 most of the hail reported was between 19 and 24 mm (dime to nickel sized). For the range from 4.0 through 4.4, there was a significant increase in the number of hail reports between 25 and 45 mm (quarter to golf ball sized). For VIL density values greater than 4.9, reported hail sizes were predominately in the range of 25 to 45 mm but with a significant number of larger reports. It is also of interest to note that for all reported hail 63 mm (tennis ball sized) and larger (six cases), VIL density was never below 4.75. Based on this information, and additional research, VIL density could become a useful tool in estimating hail size.

9. Summary

Despite the complexities associated with the NWSO Tulsa study, impressive results were obtained. A VIL density threshold of 3.5 g m^{-3} correctly identified 90% of the severe hail cases in the study and falsely identified only 2 of 36 nonsevere cases (5.5%) as severe hail producers. An unexpected result revealed VIL density might also be a useful tool for estimating expected hail size. VIL density also proved to be an effective hail indicator regardless of the season, the echo top, or the maximum VIL of a thunderstorm. Whereas VIL will always increase as thunderstorms increase in height, VIL density increases primarily due to increases in target size.

Thunderstorm motion and tilt can affect the calculation of VIL by sampling the hail shaft in more than one pixel column. The resulting "averaged" VIL results in a lower VIL density. While 3.5 g m^{-3} is a successful VIL density threshold for most thunderstorms, it was subjectively determined during the study that a VIL density value closer to 3.9 g m^{-3} may be a better threshold for slow-moving, vertical thunderstorms, where the hail shaft is more vertical and probably within the same column of pixels throughout the volume scan. Conversely, relatively fast-moving or strongly tilted thunderstorms may produce severe-criteria hail at some VIL density value below 3.5 g m^{-3}.

It should be noted that VIL density only indicates hail aloft, since the radar cannot observe hail on the ground. This can produce apparent inconsistencies in correlating VIL density to surface reports of hail. When the freezing level is quite high, and when hailstones strike significant amounts of liquid water while falling below the freezing level, melting may significantly reduce the hail size observed at the ground. This can result in reports of only small hail, or none at all, even from high VIL thunderstorms. Additional research will be required to determine the impact of high freezing levels and liquid water on hailstones.

VIL density can provide warning forecasters with a greater capability to assess the hail potential of thunderstorms. Combined with the knowledge of thunderstorm structure, VIL density should lead to improved warnings for severe hail.

REFERENCES

Greene, D. R., and R. A. Clark, 1972: Vertically integrated liquid water—A new analysis tool. *Mon. Wea. Rev.*, **100,** 548–552.

Houghton, H. G., 1985: *Physical Meteorology.* The MIT Press, 442 pp.

Paxton, C. H., and J. M. Shepherd, 1993: Radar diagnostic parameters as indicators of severe weather in central Florida. NOAA Tech. Memo. NWS SR-149, 12 pp. [Available from the National Weather Service Southern Region Headquarters, 819 Taylor Street, Room 10A26, Fort Worth, TX 76102.]

Rinehart, R. E., 1991: *Radar for Meteorologists.* University of North Dakota, 334 pp.

U.S. Department of Commerce, 1991: Doppler radar meteorological observations part c. *Federal Meteorological Handbook1 11,* 105 pp. [Available from the Office of the Federal Coordinator for Meteorological Services and Supporting Research, NOAA, USDC, Washington, DC.]