

Iwan Holleman*, Herman R. A. Wessels, Jeanette R. A. Onvlee, and Sylvia J. M. Barlag,
Royal Netherlands Meteorological Institute (KNMI), The Netherlands

1. INTRODUCTION

At the moment, a tool for the detection and display of severe weather events related to convective phenomena is being developed at the KNMI. Radar reflectivity and Doppler winds will be the primary source of information, and it will be complemented with other observations and data from numerical weather prediction (NWP) models. The KNMI operates two Gematronik C-band Doppler radars which are performing low-elevation volume scans every 5 minutes and extensive volume scans every 15 minutes. From these scans, a “constant-altitude plan-position indicator” (CAPPI) of the radar reflectivity and an echotop product, which is roughly an indicator of the cloud-top heights, are extracted. Ground clutter is removed from the CAPPI image using a statistical method (Wessels and Beekhuis, 1994). The first new product under consideration is a tool for the detection of summer hail.

Nowadays, the most direct way to distinguish between hail and rain is by using the dual-polarization radar technique which can make a direct distinction between the spherical, rotating hail stones and the non-spherical rain droplets (Smyth et al., 1999). For operational use, however, this is not feasible yet, and methods relying on single-polarization radar still have to be used. From the literature five methods for the detection of summer hail have been selected. These five different methods have been validated against hail events, that have occurred in the Netherlands during the summer of 1999.

2. METHODS

The first method which may be used to distinguish hail from severe rain is based on the CAPPI product. Because the radar reflectivity increases dramatically with increasing diameter of the scattering particles, large hail stones potentially give rise to higher reflectivities than would be possible

for rain droplets. Mason (1971) has suggested to use a reflectivity threshold of 55 dBZ for distinguishing between rain and hail.

Recently, Auer has reported on the detection of hail using a combination of the CAPPI product and cloud-top temperatures (Auer, 1994). A dependence of the optimum reflectivity threshold on the cloud-top temperature is determined. The cloud-top temperature is determined either from the infrared imagery of Meteosat or by combining the echotop product with information on the temperature profile.

The use of the entity “Vertically Integrated Liquid water” (VIL) for the detection of thunderstorms and hail has been introduced by Greene and Clark (1972). The three-dimensional radar data is converted to a plan-position indicator of the amount of liquid water present in a vertical column above a certain position. There is no agreement in literature on the best warning threshold for the detection of hail using VIL. Amburn and Wolf (1997) have proposed to use the entity “VIL-density”, defined as the ratio of the VIL value and the radar echotop height, and they suggest a universal warning threshold for the detection of hail of 3.5 g/m^3 . However, the advantage of the use of VIL-density over just VIL is disputed (Edwards and Thompson, 1998).

The current version of the “Hail Detection Algorithm” (HDA) for the WSR-88D radar network is developed by Witt et al. (1998). It is based on the hail criterium as proposed by Waldvogel et al. (1979). The method of Waldvogel for the detection of hail uses the maximum altitude at which a reflectivity of 45 dBZ is found. When this strong reflectivity extends to 1.4 km or more above the freezing level, the presence of hail is likely, and the probability of the presence of hail increases with increasing height of this reflectivity core. The height of the freezing level is determined from an NWP model.

The current HDA contains an additional algorithm that attempts to estimate the probability of severe hail (Witt et al., 1998). For this, a semi-empirical relationship between the kinetic energy flux of the hail stones and the radar reflectivity is

* *Corresponding author address:* I. Holleman, afdeling Satelliet data, KNMI, P.O. Box 201, NL-3730 AE De Bilt, The Netherlands; e-mail: holleman@knmi.nl.

used (Waldvogel et al. , 1978a,b). A “Severe Hail Index” (SHI) is calculated by vertically integrating the kinetic energy flux weighted by a reflectivity-based and a temperature-based gating function.

In this study, the performances of the CAPPI, VIL, and SHI method, the method of Auer, and the method of Waldvogel have been investigated in detail. These five detection methods are validated against on-ground hail observations, and the results are compared.

3. VERIFICATION

A systematic comparison of the output of the five selected methods for the detection of hail to on-ground observations has been conducted. Due to the small spatial extent of most hail events related to summertime thunderstorms, the 19 synops observers in the Netherlands will only report a minor fraction of the total number of hail events. Therefore, the ground truth data have been completed with hail observations by the 321 volunteers of the (rainfall) observer network of the KNMI and detailed hail damage reports from agricultural insurance companies.

The hail events can be classified using a 2-by-2 contingency table. Detected hail which is confirmed by ground truth observations will be classified as a hit (H), detected hail which is not confirmed by observations as a false alarm (F), observed hail that is not detected by radar as a miss (M), and no event as a none (N). Although much effort has been put into the collection of ground truth data, some hail events will remain unnoticed. By assuming that only a fraction η of the occurring hail events are reported, the effect of the missing ground truth data on the classification of the hail events can be investigated. In Table 1, the four classes of a modified contingency table (H' , etc.) are expressed in terms of the original classes and the fraction η . Using the modified contingency table, the apparent Probability Of Detection (POD'), False Alarm Rate (FAR'), and Critical Success Index (CSI') can be expressed in terms of the fraction η and the true POD, FAR,

Table 1: *Modified contingency table, which is valid when only a fraction η of hail events is reported.*

		Hail	No Hail
Detection:	Yes	$\eta \cdot H$	$F + (1 - \eta) \cdot H$
	No	$\eta \cdot M$	$N + (1 - \eta) \cdot M$

and POD:

$$\begin{aligned} \text{POD}' &= \frac{H'}{H' + M'} = \text{POD} \\ \text{FAR}' &= \frac{F'}{H' + F'} = (1 - \eta) + \eta \cdot \text{FAR} \\ \frac{1}{\text{CSI}'} &= \frac{H' + M' + F'}{H'} = \frac{1}{\text{CSI}} + \frac{1 - \eta}{\eta \cdot (1 - \text{FAR})} \end{aligned}$$

In this simple η -model, an incompleteness of the ground observations results in an increase of the apparent FAR' and a concomitant decrease of the apparent CSI'. A more sophisticated, two-parameter model for treating the effects of imperfect reporting on the verification of weather warnings is presented by Smith (1999).

Because most hail observations are only available per day and per municipality, the pixels in the radar images have been grouped into municipality-areas and subsequently these images have been combined to daily composites. Using a scoring program, every municipality in the Netherlands (538 municipalities, average area of 63 km²) is classified for each day as an H' , M' , F' , or N' . The final POD', FAR' and CSI' scores are obtained by combining the results for 15 days with thunderstorms in the summer of 1999.

4. RESULTS

In the assignment of a particular group of radar pixels (size: 2.4 km) having values above the warning threshold to a municipality where hail has been observed, allowance has been made for a certain spatial mismatch. A substantial gain in performance of all detection methods was observed when this spatial tolerance was increased. The method of Waldvogel, however, seems to gain the most from this increase. In contrast to, e.g., the CAPPI method (altitude 0.8 km), the method of Waldvogel uses strong radar echoes aloft (4-8 km), and therefore the horizontal spread of the on-ground hail occurrences with respect to the radar echoes is expected to be larger.

The apparent POD', FAR', and CSI' of the five selected hail detection methods have been determined as a function of their warning thresholds. The results are shown in the five sub-plots of Fig. 1. For the CAPPI, VIL, and SHI methods, warning thresholds are set on the reflectivity (in dBZ), VIL (in kg/m²), and index (in J/ms) values, respectively. For the method of Auer, the difference between the observed reflectivity and the cloud-top temperature dependent threshold is taken. Thus, a warning threshold of 0 dBZ corre-

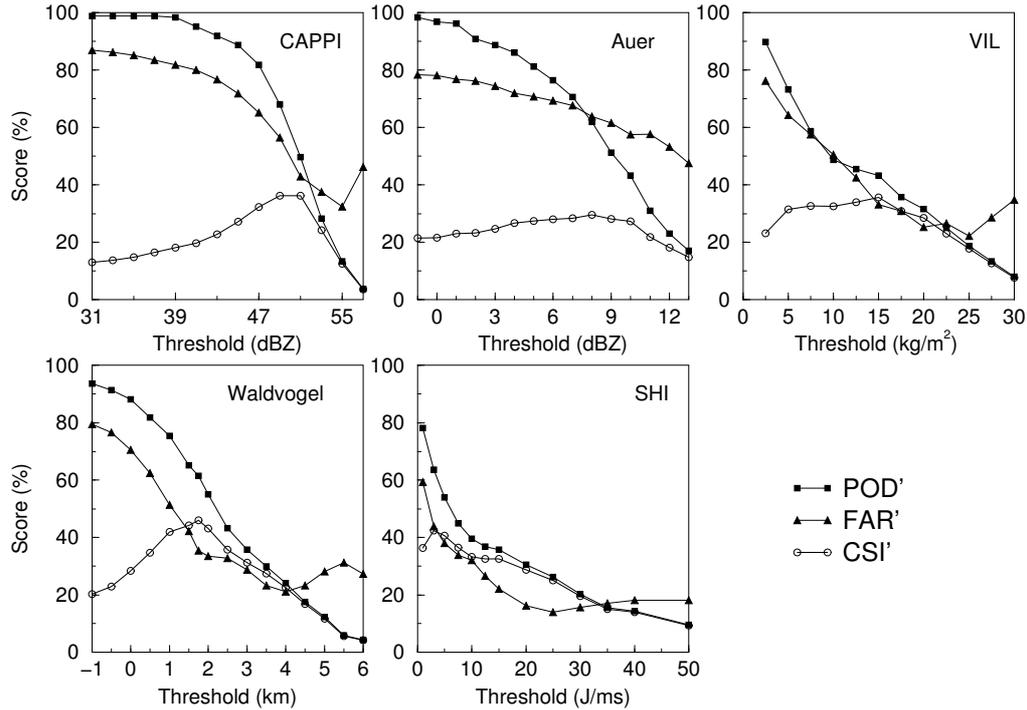


Figure 1: The scoring parameters (POD' , FAR' , and CSI') of the five different hail detection methods as a function of the warning threshold. The spatial tolerance is set at 12.5 km.

sponds to an exact reproduction of Auer’s method. The warning threshold for the method of Waldvogel is set at the difference between the maximum height of the 45 dBZ reflectivity and the height of the freezing level (in km).

Although there are large differences, the general trends in the scoring parameters as a function of the warning thresholds are rather similar for all methods. They show a decrease of both the POD' and the FAR' with increase of their warning threshold and a maximum of the CSI' at a certain threshold. The decrease of the FAR' with increase of the warning threshold, implies that, in accordance with expectations, the reliability of a detected event will increase when the warning threshold is raised.

Considering all detection methods simultaneously, a maximum POD of 99% is found, within the range of depicted warning thresholds, for the CAPPI method using low thresholds. As virtually all hail events will be accompanied by some precipitation, at most 1% of the ground truth hail reports is suspect of being inaccurate or false. The lowest FAR' is observed for the SHI method with a warning threshold of 25 J/ms, and it is about 14%. The equation for the apparent FAR' , as given previously, states that this FAR' cannot go

below $(1 - \eta)$ even when the true FAR of a detection method is 0%. Therefore, a lower limit for the fraction of reported hail events η , which is a property of the set of ground truth data only, can be determined from the lowest, apparent FAR' . Within the accuracy of the η -model and the positioning tolerance used, it is found that at least a fraction $\eta = 86\%$ of the hail events that have occurred are contained by the ground truth data used in this study.

The highest CSI' , i.e., the best performance, of 46% is observed for the method of Waldvogel using a warning threshold of 1.75 km. The method of Auer does not live up to the expectations. On our data it actually performs poorest of all methods considered. The optimum performances of the CAPPI, VIL, and SHI methods are found at lower warning thresholds than those reported in literature (Mason , 1971; Edwards and Thompson , 1998; Witt et al. , 1998). The observed discrepancies may be explained by differences in both radar calibration and climatological conditions and, for the SHI method, by the fact that it is originally designed for detection of large hail and not for detection of hail of all sizes. In addition to the highest CSI' , the largest spread in the FAR' as a function of the warning threshold is observed for the

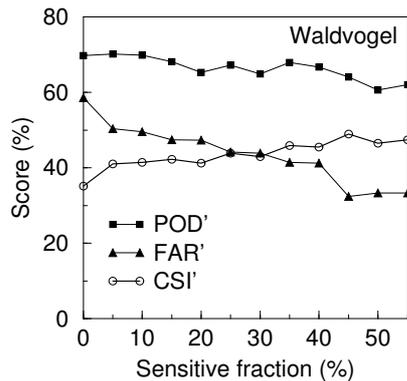


Figure 2: The scores of Waldvogel's method (threshold: 1.75 km), obtained on the ground truth data of hail damage reports only, as a function of the fraction of hail-sensitive landuse of the municipalities.

method of Waldvogel as well. This large spread in the FAR' enables the definition of several different thresholds with distinct warning properties, i.e., FAR' and resulting POD'.

The fraction of reported hail events η can be changed systematically by selecting municipalities with a certain probability of hail damage and considering the hail damage reports of the insurance companies only. Using a database of landuse in the Netherlands obtained from satellite observations between 1993 and 1995, the fraction of a municipality with hail-sensitive landuse, like crops, orchards, greenhouses, etc., is calculated. Subsequently, only the municipalities having a certain minimum fraction of hail-sensitive landuse are taken into account when the comparison of the detection methods against the hail damage reports is made. In Fig. 2 the scoring parameters of the method of Waldvogel, obtained using the optimum warning threshold of 1.75 km, are shown as function of the minimum hail-sensitive landuse fraction of the selected municipalities. In accordance with the η -model, the POD is more or less constant, the apparent FAR' is decreasing steadily, and the apparent CSI' is increasing gradually when the minimum hail-sensitive fraction is increased, i.e., when the fraction of reported hail η is increased. A maximum CSI' of roughly 49% is obtained for the method of Waldvogel in this way, which is probably close to its true CSI.

5. CONCLUSIONS

Five methods for the detection of summer hail have been compared against on-ground hail obser-

vations and reports of hail damage. Of all methods considered the one of Waldvogel scores best and that of Auer scores poorest. The CSI of 49% for the method of Waldvogel, as obtained in the present study, is somewhat higher than the CSI of 46% which can be deduced from the results presented by Waldvogel et al. (1979). The obtained CSI compares favorably to the results found for the verification of the Hail-Detection-Algorithm against events of hail larger than 6 or 13 mm in diameter (Kessinger et al., 1995). Due to the large spread in the FAR as a function of warning threshold, the warning properties of the method of Waldvogel can be altered over a wide range to fulfill the needs of different kinds of users. The method of Waldvogel will improve the detection skill significantly as compared with the CAPPI method, which is presently used at the KNMI.

REFERENCES

- Amburn, S. A., and Wolf, P. L., 1997: VIL Density as a Hail Indicator, *Wea. Forecasting*, **12**, 473–478.
- Auer Jr., A. H., 1994: Hail Recognition through the Combined Use of Radar Reflectivity and Cloud-Top Temperatures, *Mon. Wea. Rev.*, **122**, 2218–2221.
- Edwards, R. E., and Thompson, R. L., 1998: Nationwide Comparisons of Hail Size with WSR-88D Vertically Integrated Liquid Water and Derived Thermodynamic Sounding Data, *Wea. and Forecasting*, **13**, 277–285.
- Greene, D. R., and Clark, R. A., 1972: Vertically Integrated Liquid Water—A New Analysis Tool, *Mon. Wea. Rev.*, **100**, 548–552.
- Kessinger, C. J., Brandes, E. A., and Smith, J. W., 1995: A Comparison of the NEXRAD and NSSL Hail Detection Algorithms, *27th conf. on Radar Meteorology (AMS)*, 603–605.
- Mason, B. J., 1971: *The Physics of Clouds*, Clarendon Press, Oxford UK.
- Smith, P. L., 1999: Effects of Imperfect Storm Reporting on the Verification of Weather Warnings, *Bull. Amer. Meteor. Soc.*, **75**, 1099–1105.
- Smyth, T. J., Blackman, T. M., and Illingworth, A. J., 1999: Observations of oblate hail using dual polarization radar and implications for hail-detection schemes, *Q. J. R. Meteorol. Soc.*, **125**, 993–1016.
- Waldvogel, A., Schmid, W., and Federer, B., 1978a: The Kinetic Energy of Hailfalls. Part I: Hailstone Spectra, *J. Appl. Meteor.*, **17**, 515–520.
- Waldvogel, A., Federer, B., and Schmid, W., 1978b: The Kinetic Energy of Hailfalls. Part II: Radar and Hailpads, *J. Appl. Meteor.*, **17**, 1680–1693.
- Waldvogel, A., Federer, B., and Grimm, P., 1979: Criteria for the Detection of Hail Cells, *J. Appl. Meteor.*, **18**, 1521–1525.
- Wessels, H. R. A., and Beekhuis, J. H., 1994: Stepwise procedure for suppression of anomalous ground clutter, *COST-75 Seminar on Advanced Radar Systems*, **EUR 16013 EN**, 270–277.
- Witt, A., Eilts, M. D., Stumpf, G. J., Johnson, J. T., Mitchell, E. D., and Thomas, K. W., 1998: An Enhanced Hail Detection Algorithm for the WSR-88D, *Wea. Forecasting*, **13**, 286–303.