The NSSL Hydrometeor Classification Algorithm in Winter Surface Precipitation: Evaluation and Future Development

Kimberly L. Elmore*
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

(Manuscript received 23 August 2010, in final form 7 March 2011)

ABSTRACT

The National Severe Storms Laboratory (NSSL) has developed a hydrometeor classification algorithm (HCA) for use with the polarimetric upgrade of the current Weather Surveillance Radar-1988 Doppler (WSR-88D) network. The algorithm was developed specifically for warm-season convection, but it will run regardless of season, and so its performance on surface precipitation type during winter events is examined here. The HCA output is compared with collocated (in time and space) observations of precipitation type provided by the public. The Peirce skill score (PSS) shows that the NSSL HCA applied to winter surface precipitation displays little skill, with a PSS of only 0.115. Further analysis indicates that HCA failures are strongly linked to the inability of HCA to accommodate refreezing below the first freezing level and to errors in the melting-level detection algorithm. Entrants in the 2009 American Meteorological Society second annual artificial intelligence competition developed classification methods that yield a PSS of 0.35 using a subset of available radar data merged with limited environmental data. Thus, when polarimetric radar data and environmental data are appropriately combined, more information about winter surface precipitation type is available than from either data source alone.

1. Introduction

The Weather Service Radar-1988 Doppler (WSR-88D) network will soon undergo an upgrade that will provide all radars with polarimetric capability (Ryzhkov et al. 2005b). Among the expected benefits of this upgrade is the ability to use polarimetric information to help to distinguish hydrometeor type (e.g., Herzegh and Jameson 1992; Zrnić and Ryzhkov 1999; Vivekanandan et al. 1999; Ryzhkov et al. 2005b). The National Severe Storms Laboratory (NSSL) has developed a hydrometeor classification algorithm (HCA; Zrnić et al. 2001; Straka et al. 2000; Park et al. 2009) for use in warm-season convective weather; it, along with a melting-layer detection algorithm (MDA) and a rainfall-estimation algorithm, is the only such algorithm currently scheduled for the initial deployment of the dual-polarization (dual pol) WSR-88D network. Yet, a top-ranked expectation recently expressed by the operational forecast community is the ability to “determine the precipitation type during winter and delineate where rain/snow transition lines likely exist, particularly in areas where spotters or surface observations are sparse...” (emphasis added, Cocks 2010). The HCA’s performance for winter surface precipitation has not been quantitatively evaluated, however.

Because the NSSL HCA was developed primarily for use in deep, moist convective storms, it is useful to determine its applicability in cold-season regimes. To that end, this paper examines the NSSL HCA performance in winter precipitation at (or very near) the surface on the basis of surface-based observations of precipitation type.

2. Experimental approach and data

a. Surface observations

To verify quantitatively the HCA performance for precipitation type at the surface requires both radar and direct observations over a large area and at frequent time intervals throughout precipitation events. Placing enough scientists in the field, or even providing paid
observers, during an entire season solely for precipitation-type observations is not only wasteful but beyond the means of any field project. Upon careful consideration, none of the existing observational networks suffice for this work. Instead, voluntary observations within 150 km of the Norman, Oklahoma, radar (KOUN; Fig. 1) are solicited from the general public through public media announcements and a Worldwide Web interface. No specific requirements or affiliations are imposed upon participants, and no identifying information is retained. The project that collected the data presented here is called the Winter Precipitation Identification near the Ground project (W-PING; see online at http://www.nssl.noaa.gov/projects/winter/form.php).

The W-PING project was launched on the eve of a well-forecast, heavily publicized winter storm in late November 2006, and the public response was exceptionally high. The winter of 2006/07 was fortunately active, with three significant ice and snow events in and around the KOUN dual-pol prototype radar located at the University of Oklahoma Westheimer Airport in Norman. The data presented here come from events on 29–30 November 2006 and 11–14 and 19–20 January 2007. There are other data-collection projects that utilize the public [e.g., the National Oceanic and Atmospheric Administration (NOAA) Storm Data publication], but the W-PING data are particularly attractive because they are coordinated with polarimetric radar data collection. Although only three major events are included in these data, the range of precipitation types associated with these three events runs the entire gamut of winter precipitation types that can be reported.

The W-PING project Web site contains material about how to determine whether precipitation is drizzle, rain, sleet, or snow and discusses freezing precipitation. Through this Web site, the public are invited to report the primary precipitation type observed during winter storms using the following categories: none, rain, freezing rain, drizzle, freezing drizzle, snow, ice pellets/sleet, graupel/snow grains, and hail. Implicit with any such project is the assumption that the public can reliably distinguish the difference among, say, rain, snow, and sleet. Although errors undoubtedly exist within the verification data, these errors are assumed to be random and unbiased.

Public reports are entered through a Web-based form, selecting the primary precipitation constituent (the one noted by the observer as the most prevalent) using radio buttons, along with the time of the observation and the location (in latitude and longitude) of the observation. If observers do not know their latitude and longitude, they determine it using links to geolocation Web sites, which return a reasonably accurate latitude and longitude given an address. Data are quality checked, primarily on the basis of temporal and spatial consistency. If, for example, all observations within a specific time and area report snow, a single observation of hail in their midst is doubtful and is therefore removed. Observations associated with obvious time and location errors are also removed.

Even though the W-PING project report form provides many categories, they are not aligned with the categories produced by the HCA, which are no echo, light/moderate rain, heavy rain, rain/hail mix, big drops, dry snow, wet snow, graupel, crystals, anomalous propagation/ground clutter, and biological targets. It can be argued that even trained observers would have difficulty with some of the HCA categories and therefore that expecting the public to provide observations over these same categories would be unreasonable. Even so, the public are assumed to be capable of distinguishing frozen from liquid precipitation and certainly of distinguishing rain from snow.

The misalignment of categories is managed by collapsing all categories for both observations and HCA types into only three: liquid precipitation, frozen precipitation, and
Figure 2 lists the raw confusion matrix for the W-PING-observed types and HCA classes. Any rain or drizzle, whether heavy, moderate, or light, freezing or not, is liquid. Any kind of frozen water substance, such as ice crystals, graupel, ice pellets, wet snow, dry snow, or hail, is classed as frozen. During the event, the public are specifically encouraged to enter observations of “none” if they observe no precipitation falling.

Within these data, there are 356 observations of liquid precipitation (29.4% of the total observations), 155 observations of no precipitation (12.8% of the total) and 699 observations of frozen precipitation (33.0% of the total). Within the frozen category, 538 are sleet/graupel (77%) and 165 are snow (23.6%). A single report of hail exists but is not used.

That this study uses public observations does not mean that public observations can, or necessarily should, completely replace automated observing systems for similar studies. The primary automated observing system in use within 150 km of the KOUN radar is the Automated Surface Observing System (ASOS), in which the light-emitting diode weather indicator (LEDWI) system is primarily responsible for determining precipitation type. However, the LEDWI has known difficulties with some precipitation types, such as drizzle and freezing drizzle (Wade 2003) and cannot report the presence of ice pellets (National Oceanic and Atmospheric Administration 1998). Public observers, on the other hand, can reliably identify ice pellets. In addition, public observers are more widespread than the ASOS stations. Therefore, this investigation uses public reports as the sole source of observed precipitation type.

| TABLE 1. W-PING vs HCA results for the 2006/07 season. |
|-------------|----------|-----------|----------|----------|
|             | $\sigma_{\text{liquid}}$ | $\sigma_{\text{none}}$ | $\sigma_{\text{frozen}}$ |
| $f_{\text{liquid}}$ | 233      | 30        | 530      |
| $f_{\text{none}}$   | 119      | 124       | 151      |
| $f_{\text{frozen}}$ | 4        | 1         | 18       |

Fig. 2. Raw confusion matrix for W-PING and HCA categories. Numbers represent how well classes form each system “match.” Observed hydrometeor type is represented by columns, and HCA-derived types are represented by rows. Thus, the value at the intersection of the column labeled “snow” and the row labeled “Lt Rain” yields how many times observers reported snow when the HCA was generating a classification of light rain. The light gray area at the upper left shows how categories are collapsed into “none,” the clear areas show how categories are collapsed into “liquid,” and the darker gray areas along the right and lower third of the table show how categories are collapsed into “frozen.”
b. Radar data

Data fields produced by KOUN and archived for later analysis are mean radial Doppler velocity \( V_r \), reflectivity \( Z \), spectrum width \( SW \), differential reflectivity \( Z_{DR} \), differential phase \( \Phi_{DP} \), specific differential phase \( K_{DP} \), and cross-correlation coefficient \( \rho_{HV} \). Only \( Z_{DR} \), \( \Phi_{DP} \), \( K_{DP} \), and \( \rho_{HV} \) are polarimetric variables. The two variables \( SW \) and \( \Phi_{DP} \) are not used in any analyses presented here. KOUN radar data are calibrated in house at the NSSL. Because the signal-to-noise ratio in winter precipitation is considerably less than in warm-season precipitation, noise corrections for \( \rho_{HV} \) and \( Z_{DR} \) are also performed (Ryzhkov et al. 2005a; Bringi and Chandrasekar 2001).

Each ground observation is associated with radar data that are both high enough above the ground to avoid serious ground-clutter contamination and low enough to the ground to sample a volume that is likely to contain what an observer on the ground will see. A single-elevation-angle rule is insufficient in this regard, and therefore radar data closest in time and space to the observation and centered between 300 and 1200 m above ground level are used here. Within the 300–1200-m layer, data from the lowest elevation scan are chosen so as to avoid multiple radar data values for a single ground observation. Within this layer, near-surface thermodynamic profiles will lead to a mixture of hydrometeor type within the radar sample volume. This altitude interval is a compromise between providing sufficient data for analysis and constraints to minimize problems inherent in radar sampling.

Radar data consist of sample volumes, each 250 m in radial extent and 1° in azimuth extent (1° is the beam spacing). Each sample volume contains radar values and a hydrometeor classification. Radar data values extracted over each observation consist of the simple mean taken over a 5×5 (range pulse volumes × azimuth beams) window or stencil centered over the observation (times sign, or X, in Fig. 3). Because a mean class has no meaning, the hydrometeor class associated with each ground observation is defined as the most prevalent class in the 5×5 stencil centered over each observation. Observations associated with a resulting HCA classification of anomalous propagation/ground clutter are removed. Once these quality checks are complete, HCA precipitation types are compared with the reported precipitation types.

For the three main events, about 2650 logged observations exist. After quality-control processing, about 2500 remain. Of these 2500, 1210 satisfy the criteria stated above. The remaining ~1300 observations are too far away for the lowest scan to reside below 1200 m. Some other minor events occurred in 2006/07 but are not used here.

3. Results and analysis

a. Skill score

A generic 3×3 confusion matrix (Table 1) is arrayed such that rows are HCA classes and columns are observations. The first row is for HCA classifications of liquid, the second row is for HCA classifications of none, and the third row is for HCA classifications of frozen precipitation. Likewise, the first column is for observations of liquid precipitation, the second column is for observations of no precipitation, and the third column is for observations of frozen precipitation. Values that appear along the main diagonal represent the number of classifications that agree with corresponding observations. Any off-diagonal values represent misclassifications. For example, the cell at the bottom of the first column represents the number of cases misclassified as frozen but for which observations indicate liquid precipitation is present.

The Peirce skill score (PSS; Peirce 1884; Jolliffe and Stephenson 2003; Wilks 2006) is the evaluation statistic used here. For multiple classes (two or more), the PSS is given by

\[
PSS = \frac{\sum_{i=1}^{J} p(f_i,o_j) - \sum_{i=1}^{J} p(f_i) p(o_j)}{1 - \sum_{j=1}^{J} [p(o_j)]^2},
\]

where \( I = J =  \) number of categories. The PSS ranges from −1 to 1, with 0 indicating no skill beyond guessing based upon the climatological values within the dataset. For example, the observations contain 29.3% liquid,

![Fig. 3. Values at the times sign (×) are computed from the 5×5 stencil of radar pulse volumes. For numerical values, the value at × is a simple mean. For the HCA class, the class at × is the most prevalent class within the 5×5 region.](https://example.com/fig3.png)
12.8% none, and 57.9% frozen. Thus, random guessing by generating these percentages for each category will result in an expected score of zero. The PSS is also equitable (Gandin and Murphy 1992) and so is resistant to hedging or gaming. The confusion matrix for the 1210 HCA classifications matched to observations is given in Table 1. In the ideal case, the HCA should generate a type that tends to match what an observer sees, yielding a PSS near 1. The actual HCA score is shown in Fig. 4, with the 95% empirical confidence interval derived from bootstrap resampling.

b. Specific scores

When applied to winter precipitation, HCA appears to classify too many cases as liquid when either frozen or none is reported (Table 1). In fact, it seldom classifies anything as frozen. Although 57.9% of the observations report frozen precipitation, only 1.8% of the classifications from the HCA are for frozen.

When HCA is used simply to determine the presence of frozen precipitation (“frozen” or “something else”), the PSS decreases dramatically [Fig. 4; “HCA (fzn)”].

Although this score is statistically different from zero, in practice such performance is probably imperceptible from random guessing. In a similar way, if HCA is used to determine the presence or lack of only liquid precipitation, the PSS drops to a value that is statistically no better than zero [Fig. 4; “HCA (liq)”].

There unfortunately is no ideal way to generate scores for HCA regarding liquid versus frozen because in some cases HCA generates a liquid or frozen classification when the observation is none, and vice versa. A naïve approach to HCA performance for only liquid versus frozen classification is simply to extract the four corner values of Table 2. Such an approach is purely ad hoc and uses only 788 of the 1210 available cases. Yet, doing so results in essentially zero skill as measured by the PSS [Fig. 4; “HCA (liq vs. fzn)”]. The HCA does much better when used to classify whether precipitation of any kind is present (“precipitation” versus “no precipitation”), shown by the score “HCA (any)” in Fig. 4. It is clear that most of the skill for a three-category HCA comes from its ability to detect precipitation or no precipitation.

That HCA is unable to determine perfectly the presence or absence of observed precipitation is evidence that the mere presence of a radar echo does not equate to precipitation at the surface. There are cases in which an echo is detected but no precipitation is observed at the ground (false alarms), and there are other cases in which an observer detects precipitation that is not detected by the radar (missed detections). No precipitation intensity estimates are requested of observers, and therefore very light precipitation may be detected at the surface but be too light to create a sufficient radar return (missed detection). Or, precipitation may not be falling at the observation point yet may be falling elsewhere nearby and so be contained within the radar sample volume (false alarm). Ground clutter may be misclassified as a precipitation type (false alarm), and so on. The mere presence of a radar return is insufficient for declaring the presence of precipitation, an issue that HCA is intended to help to address.

c. Explanations of poor performance

The poor performance of the HCA on frozen precipitation may be explained in part because frozen precipitation is sometimes mixed with varying amounts

---

**Table 2. W-PING vs HCA results for the 2006/07 season after all observations of sleet are removed.**

<table>
<thead>
<tr>
<th></th>
<th>(o_{\text{liquid}})</th>
<th>(o_{\text{none}})</th>
<th>(o_{\text{frozen}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_{\text{liquid}})</td>
<td>233</td>
<td>30</td>
<td>105</td>
</tr>
<tr>
<td>(f_{\text{none}})</td>
<td>119</td>
<td>124</td>
<td>47</td>
</tr>
<tr>
<td>(f_{\text{frozen}})</td>
<td>4</td>
<td>1</td>
<td>14</td>
</tr>
</tbody>
</table>
of rain. Observers are not asked to report mixed precipitation because doing so reliably requires unavailable training. Another likely contributor is that the current HCA does not account for refreezing below the melting level, a common occurrence in winter precipitation (Stewart 1985). Raindrops can refreeze very close to the surface, well below the radar beam (save for very close ranges), and so are not observed by the radar. In such cases, any variant of a radar-only HCA will certainly fail when liquid precipitation refreezes. To capture refreezing requires vertical temperature information near the surface such that ice pellet formation can be inferred. There are some legitimate observations of graupel during these storms but the public may not be skillful in applying that classification. Thus, all observations of graupel are treated as ice pellets/sleet.

If ice pellets are the primary source of low HCA skill, a reasonable test is to remove cases associated with reports of either sleet or graupel (the latter is removed because the public probably cannot reliably distinguish between sleet and graupel) and to score what is left. The underlying hypothesis is that removing cases in which HCA is guaranteed to fail should yield an improved skill score. There are 533 observations of sleet and/or graupel within the database, leaving 677 cases for analysis (Fig. 2). Table 2 shows the results after all reports that contain observations of refrozen precipitation are removed. The resulting PSS improves markedly ["HCA (no sleet)" in Fig. 4]. On the basis of a permutation test (Efron and Tibshirani 1993), this improvement is statistically significant in comparison with the HCA score that includes sleet, with a significance value of <0.001. Hence, the inability of HCA to accommodate refreezing of hydrometeors below the first freezing level is a significant error source within the current HCA when applied to winter precipitation.

HCA also does poorly with snow (Fig. 2). This is due, in part, to erroneous information from the MDA. Melting layers from the MDA are used in lieu of any environmental data (such as a sounding) as a “first guess.” Within Fig. 2 are 165 observations of snow, of which 104 are classified by HCA as liquid. In every such case, this misclassification occurs because the MDA has erroneously identified a melting layer. Because no frozen precipitation can exist below the melting layer within the current HCA, a classification of snow cannot occur.

4. Statistical classifiers

a. Some “toy” examples

Given these data, exploring how effective polarimetric variables are in discriminating among the three basic classes (frozen, none, and liquid) in light of additional environmental variables seems reasonable. As a statistical exercise, the data are randomly divided into training and testing sets consisting of 847 and 363 cases, respectively. Applying a linear discriminant analysis (LDA; Fisher 1936; Wilks 2006) to only the radar parameters within the training set \( Z, Z_{DR}, R_{HV}, \) and \( K_{DP} \) and then applying the discriminant function to the testing set results in a PSS of 0.138 ("Radar LDA" in Fig. 5, second from left). On the basis of a permutation test, this PSS is statistically indistinguishable from the base HCA results that include refrozen precipitation. Hence, HCA is no better (or worse) than a simple linear discriminant function applied to only the radar parameters. A quadratic discriminant analysis (QDA; Wilks 2006) function similarly applied performs in a statistically similar manner ("Radar QDA" in Fig. 5, third from left). Thus, in both first- and second-order polynomial models, polarimetric parameters alone are ambiguous.

As a next step, the observations are associated with environmental data derived from the Rapid Update Cycle (Bleck and Benjamin 1993; see online at http://www.ruc.noaa.gov) Model. The environmental data include the \( u \) and \( v \) wind components, temperature, and relative humidity, all at 2 m, plus the height of the lowest freezing level. No other environmental parameters are extracted.

When an LDA is developed using only the environmental parameters without any radar data fields and then is applied to the testing data, the score denoted as “Env LDA” in Fig. 5 results (fourth from left). A
permeability is shown at “Env + Z QDA” in Fig. 5 (second from right). The best QDA uses all of the environmental variables plus the $\rho_{HV}$ and $K_{DP}$ radar fields and is shown as “Best QDA” in Fig. 5 (rightmost). This combination shows an improvement in classification skill over both the environmental-only QDA and the QDA that uses both the environmental parameters and reflectivity. It is curious that adding reflectivity to the QDA that uses all environmental parameters, $\rho_{HV}$, and $K_{DP}$ results in reduced skill, although the reduction is not statistically significant. Hence, in both linear and quadratic (non-linear) space, polarimetric radar variables enhance skill. That the best QDA is more skillful than the best LDA suggests that a nonlinear classifier will probably exhibit better overall performance.

b. More-sophisticated techniques

The preceding analysis is purely exploratory and is statistically driven. There is no reason to expect that physical processes are best described by simple first- or second-order polynomials or effects. Yet, both are clearly important for classifying winter hydrometeor types into the three basic categories. The environmental data are purposely incomplete and so provide a limited description of the environment. Thus, these discriminant functions are not further explored or interpreted because the point of the exercise is made: winter surface hydrometeor classification can be substantially improved beyond that of the current HCA when nonlinear associations between both the environment and the radar parameters, driven by observations, are included. It is clear that when polarimetric radar data are combined with environmental data better classification performance is realized than when using either radar or environmental data alone, and the performance significantly improves over the current HCA.

With the above results in mind, these data were made available to contestants as part of the American Meteorological Society (AMS) second annual artificial intelligence (AI) competition (Elmore and Richman 2009). As described above, contestants use the training set to generate a classification technique and, then, apply it to the testing set. Although not the original intent, retrospectively the AI competition proves an expeditious way to investigate the feasibility of improving HCA based on efforts of some of the finest minds available using whatever means they deem appropriate.

The AI competition results are shown in Fig. 6, including those from the HCA analysis, the best LDA, and the best QDA described above. Using a permutation test, the statistical difference between the LDA and the first-place scheme appears at the $p = 0.049$ level. In strict terms, the scores for these two techniques differ at the 5% level. Although there are differences between entries, even the lowest-scoring technique—the LDA—is considerably (and statistically) better than the HCA.

Although the methods employed to achieve these improved results in the AMS AI competition are varied, all of the various entries utilized some type of observation-driven, statistically principled technique. The various techniques used included a support vector machine (Sullivan 2009); a neural net–decision tree combination (Alam and Goossensaert 2009); a neural net–random forest combination (Pocernich 2009); a three-model approach of nearest neighbor, pure neural net, and simple decision tree (Lakshmanan 2009); an annealing and gradient descent (Gordon, et al. 2009); a variant of k-nearest neighbor classifier (McCandless 2009); and a different neural-net variant (Pelliccioni et al. 2009). Of all of these techniques, Sullivan (2009), Alam and Goossaert (2009), and Pocernich (2009) submitted the top three methods. These results demonstrate that methods based on observation-driven, statistically principled approaches provide substantive (rather than merely statistical) improvements over the current HCA for winter surface precipitation, even when there are limitations to both the environmental and the radar data.

5. Conclusions and future work

In its current form, the NSSL HCA is not suitable for determining hydrometeor type near the ground in winter precipitation. This result is not surprising because the current HCA does not accommodate refreezing of hydrometeors. Any method that cannot accommodate refreezing of hydrometeors is bound to fail in such relatively common circumstances. In addition, HCA depends almost exclusively on radar data that, because of radar sampling geometry, degrade with increasing range from the radar. To be fair, the NSSL HCA was never intended for classifying winter precipitation type at the surface, but the temptation to misuse it in such a manner remains. Should a user yield to that temptation, the
results will be disappointing. Even so, the disappointing HCA performance in winter precipitation is a shortcoming of the algorithm and not of the polarimetric radar data. Even for a simple quadratic discriminator, polarimetric data clearly add information and so contribute to skill.

No doubt remains that a major contributor to HCA misclassifications lies in the current algorithm’s inability to accommodate refreezing of hydrometeors, along with erroneously identified melting levels. Suggestions for improvement that utilize some environmental information appear in Park et al. (2009), but these modifications have yet to be either thoroughly tested or implemented. The results that appear here address HCA performance for determining winter surface precipitation type for the HCA that will be fielded with the WSR-88D polarimetric upgrade.

The AI competition techniques use observations to drive algorithms that all utilize both environmental and polarimetric radar data nonlinearly and show that the polarimetric radar data is an integral component of any winter surface HCA (WSHCA). One of the most interesting results from the AI competition is that there appears to be no statistically significant difference among any of the methods employed by the entrants when measured with the PSS. Of all entrants, only the difference between the Pelliccioni et al.’s (2009) and Sullivan’s (2009; the first-place entrant) PSS scores comes closest to being significant at the 5% level. Except for the Sullivan (2009) entry, there is also no statistical difference between either the LDA or the Pelliccioni et al. (2009) entry and any of the other entrants. This result suggests that there may be little perceived difference between any of the top 10 methods were they to be employed and that no particular method clearly outshines the others. This characteristic uncertainty is probably driven by withholding information about both the vertical structure of the environment and the vertical structure of the polarimetric radar variables. The NSSL HCA PSS is statistically much worse than any other method employed, however, which points to the value of environmental information as a part of any future WSHCA development.
The data used for the competition represent a subset of the total currently available. For example, no vertical profiles of either radar data or environmental data are used. There is little doubt that a winter surface HCA can be developed based on a careful choice of additional variables, both radar and environmental, driven by knowledge of physical precipitation processes and, in particular, observations of precipitation type. Such a WSHCA may reasonably be expected to yield significant improvements in skill beyond that shown by the AI competition and so is considerably better than the current HCA. A WSHCA could be developed that provides probabilistic classifications, which could provide useful guidance to operational meteorologists. Probabilistic output is not likely to be part of the initial WSHCA development, however. On the basis of the current HCA performance in winter surface precipitation, development of a WSHCA is not only clearly imperative but is, in fact, under way at the NSSL.

To proceed with such development, and for the resulting algorithm to have meaningful, quantifiable skill and accuracy, requires, first and foremost, many thousands of surface observations of precipitation type. Because the public can clearly discern the difference between frozen and liquid precipitation (e.g., the difference between sleet and rain), the easiest way to collect these data is through public observations during winter events that are also observed by polarimetric radar. If ice pellets and freezing drizzle are considered, only human observations will suffice because the ASOS LEDWI cannot properly identify ice pellets and has trouble with drizzle. Collecting these data from the public is fortunately inexpensive and, as a bonus, allows the public to become involved in science in a substantive, meaningful way. Experience shows that quality-control efforts are needed to ensure that the data provided by the public are consistent. Such a project is already under way and will be expanded as polarimetric capability is added to the existing WSR-88D network.

Acknowledgments. John Krause provided the HCA classifications based on the most recent algorithm and provided invaluable insight about the inner workings of the NSSL HCA. Vicki Farmer developed the Worldwide Web interface used for collecting the public observations. Terry Schuur generously gave his time and expertise with patient explanations and diligent calibration of KOUN data. Alexander Ryzhkov and Dusan Zrnić patiently answered questions and made many useful comments. Heather Reeves and V. Lakshmanan provided very helpful comments. Paul Schlatter and two anonymous reviewers provided valuable suggestions for improvement. This work was supported in part by the High Performance Computing and Communications Office of the National Oceanic and Atmospheric Administration, and also by the FAA Aviation Weather Research Program through Interagency Agreement DTFAWA-08-Z-80002.

REFERENCES


Pelliccioni, A., R. Contraceo, and F. Pungi, 2009: Application of neural net models to classify and to forecast the observed precipitation type at the ground using the artificial intelligence competition data set. Preprints, Seventh Conf. on Artificial Intelligence Applications to Environmental Science, Phoenix, AZ, Amer. Meteor. Soc., 3.3. [Available online at http://ams.confex.com/ams/pdfpapers/152329.pdf.]


