

11R.4 CLASSIFICATION OF HYDROMETEORS AND NON-HYDROMETEORS USING POLARIMETRIC C-BAND RADAR

Jonathan J. Gourley, Pierre Tabary, and Jacques Parent du Chatelet

Meteo-France, DSO, Centre de Meteorologie Radar, Trappes, France

1. INTRODUCTION

Several European operational weather agencies are considering upgrading their C-band radar networks with polarimetric capabilities based on the prospects of getting improved quantitative precipitation estimates and the ability to discriminate hydrometeors. Polarimetric variables are sensitive to particle concentrations, shapes, sizes, and orientations. In addition, recent studies (e.g., Zrníc et al. 2005) have shown that common contaminants to traditional Z_H products can be significantly mitigated through the use of multiparameter observations. This study presents a methodology to readily adapt and tune fuzzy logic algorithms using polarimetric observations at C-band. Because the design of the weighting functions is based entirely on observations, this method may be applied in the tropics or mid-latitudes at X-, C-, or S-band wavelengths. The empirical approach of computing membership functions is presented. The algorithm is then applied to observations collected by the Meteo-France C-band polarimetric radar with the primary intention of discriminating between hydrometeors and non-hydrometeors.

Typically, data quality algorithms are formulated in a fuzzy logic sense. Fuzzy logic algorithms differ from decision tree logic in that they enable the incorporation of observations from multiple parameters that may have non-mutually exclusive conditions describing a given hydrometeor type. They rely on weighting functions that are smooth, may overlap, and can be multidimensional as in Straka et al. (2000). The shape of these weighting or membership functions controls the behavior and thus performance of the precipitation-typing algorithm. The choice of these functions is subjective varying from Beta functions (Liu and Chandrasekar 2001) to more simple trapezoidal functions. The ranges of values they cover are also based on subjective knowledge, observations, or simulations. An alternative approach is adopted in this study where the membership functions are derived directly from observations using Gaussian kernel density estimation (Silverman 1986). The technique requires an analyst to restrict the spatial and temporal domains so that a desired hydrometeor type is isolated. Membership functions are then derived automatically. This empirical approach is believed to be advantageous for the following reasons:

- membership functions are specific to the radar wavelength, beam propagation characteristics at radar location, typical scatterers (e.g., biological) encountered in regime, scanning strategy, and volume coverage pattern
- functions may be updated and tuned as new observations of weather phenomena become available
- methodology is easily extended to accommodate multidimensional functions
- analysis of functions leads to a posteriori understanding of backscatter and propagation properties of hydrometeors

2. METHODOLOGY

a. Background

The Meteo-France C-band polarimetric radar, located approximately 30 km to the southwest of Paris, has been collecting observations in an operational setting for over a year. The first task related to hydrometeor identification is the simple partitioning between hydrometeors and non-hydrometeors. The methodology developed herein will ultimately be used for discriminating between different hydrometeor species (e.g., rain, melting snow, hail), but this initial phase of work is aimed at improving data quality. The choice of crisp inputs (input variables) and crisp outputs (particle classes) is thus simplified for a limited application. Moreover, the need for multi-dimensional membership functions is not necessary for this study, but the extension of the methodology to accommodate these functions is straightforward.

b. Definition of crisp inputs (variables)

The Trappes radar is equipped with linear polarization capabilities in that it transmits horizontally and vertically polarized waves. The two receiving channels, which have nearly identical waveguide runs, operate in parallel and thus enable the simultaneous reception of polarized signals. The relevant variables collected are reflectivity at horizontal polarization (Z_H), differential reflectivity (Z_{DR}), correlation coefficient between horizontally and vertically polarized waves at 0 time lag ($\rho_{HV}(0)$), and differential propagation phase (Φ_{DP}). A detailed study has been undertaken in order to identify, quantify, and correct for the effects of noise,

* Corresponding author address: Jonathan J. Gourley, Centre de Meteorologie Radar, DSO, Meteo France, 7 rue Teisserenc-de-Bort 78195, Trappes, France; email: jonathan.gourley@meteo.fr

miscalibration, near-radome interference, and system offset in initial Φ_{DP} measurements (Gourley et al. 2005a). In addition, an attenuation correction procedure has been developed and implemented for the known reductions in Z_H and Z_{DR} at C-band frequencies (Gourley et al. 2005b). It was discovered that the textures of several variables were quite different, especially for discriminating between hydrometeors and non-hydrometeors. For this reason, the root mean squared error of all polarimetric variables is computed in a 3x3 pixel window (corresponding to 1.5°x720m). Lastly, Doppler radial velocities (v) and temperatures (T) are available for detailed particle-typing studies.

All variables and their derivatives are stored on a grid with a resolution of 240m x 0.5°. While the radar collects data up to an elevation angle of 90°, this data quality study primarily deals with observations at an elevation angle of 1.5°. Data from low elevation scans are available every 5 min.

c. Identification of hydrometeor and non-hydrometeor classes

The relevant hydrometeor and non-hydrometeor classes for this study are precipitation (PPT), echoes from ground clutter and anomalous propagation (AP/GC), chaff echoes (CHAFF), and scatterers in clear air (CLEARAIR). The latter three classes are contaminants to Z_H products. AP/GC is a commonly observed phenomenon that results from power returns from natural and anthropogenic features typically situated near the radar. Chaff, or synthetic fibers released from military aircraft, isn't observed as frequently but can take the appearance as a fine line or even as an intense cell on Z_H images. Echoes from clear air are not as well understood, but may be a result of biological targets.

An expert analyst identifies a case where a given class is obvious and prevalent. Additional sources of information are utilized in order to unambiguously identify the class of interest. For example, in order to isolate AP/GC a case is chosen on a sunny day where there are no precipitating echoes in the vicinity. Next, the analyst defines the azimuths, ranges, and altitudes at which the phenomenon is occurring. For AP/GC, the search domain would be restricted to elevation angles of 0.4° near the radar (i.e., ranges < 10 km) at all azimuths. It is important to note that no thresholding on polarimetric variables is used in this step. This enables the derived membership functions to remain unconditional and objective.

d. Gaussian kernel density estimation

The distribution of each distinct input is computed for the hydrometeor type using Gaussian kernel density estimation (Silverman 1986). The kernel density estimate is defined as:

$$\hat{f}(x) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{i=1}^n e^{-\left[\frac{1}{2}\left(\frac{X_i-x}{\sigma}\right)^2\right]} \quad (1),$$

where $\hat{f}(x)$ is an estimate of the data density, σ is the smoothing parameter or bandwidth, n is the total number of data points, and X_i is the i^{th} value. The bandwidth controls the number of bumps or the smoothness in the estimation of $f(x)$. This parameter is estimated using the so-called Silverman's rule ($\sigma=1.06*\text{stdev}*n^{1/5}$). In effect, a Gaussian curve is produced at each observation along the abscissa. The width of these curves is determined by the value of σ . Next, linear superposition is used to sum up all the curves in order to arrive at a continuous estimate of the data density. This method has been extended to the bivariate or multivariate case by using a multiplicative kernel.

Examples of 1-D membership functions that are particularly useful in segregating hydrometeors from non-hydrometeors are shown in Figs. 1 and 2. The distribution of $\rho_{HV}(0)$ for PPT is relatively narrow near large values (mode=0.97). CHAFF and CLEARAIR may be adequately segregated from PPT using this parameter alone, but it is noted that there is substantial overlap between the PPT and AP/GC distributions. Additional parameters are thus needed in order to successfully identify scatterers caused by precipitation. Observations of individual polarimetric images indicated that non-weather phenomena have much greater spatial heterogeneity over small scales. To the human eye, images of Φ_{DP} , for example, appear noisy and granular. The $\text{text}(\Phi_{DP})$ conveys this information to the algorithm as shown in Fig. 2. In this case, the distributions of PPT and AP/GC are well separated. The PPT distribution is much more narrow with a peak near low values of $\text{text}(\Phi_{DP})$. The combination of the membership

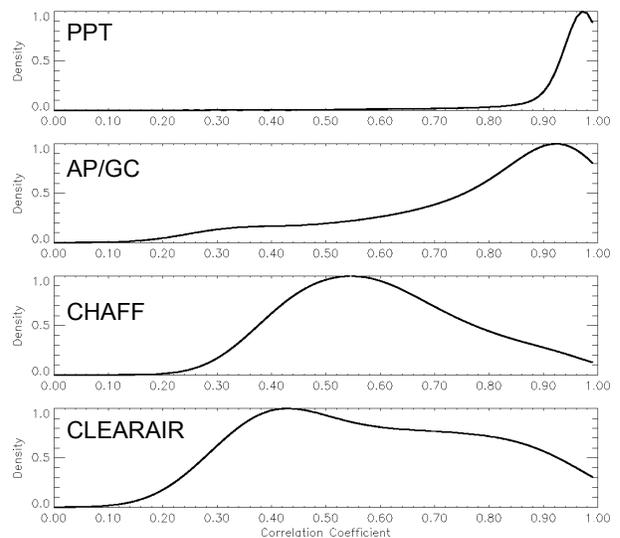


FIG. 1. Empirically-derived membership functions for $\rho_{HV}(0)$.

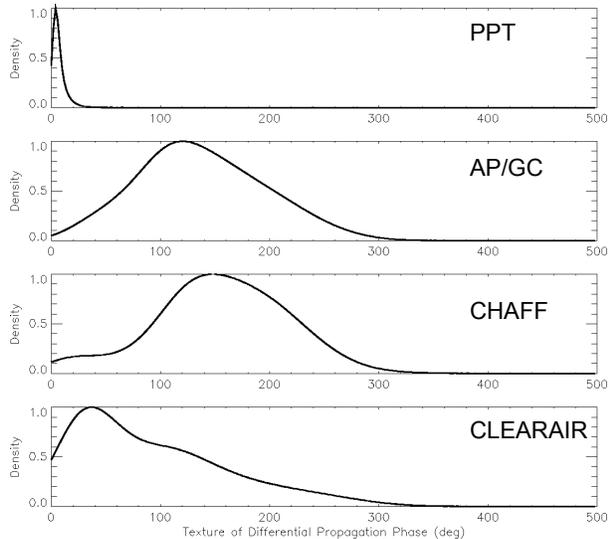


FIG. 2. Empirically-derived membership functions for texture of Φ_{DP} .

functions shown in Figs. 1 and 2 in addition to $text(Z_{DR})$ and $text(\rho_{HV}(0))$ are used in the fuzzy logic algorithm for hydrometeor vs. non-hydrometeor discrimination. These variables were chosen subjectively through experimentation.

e. Fuzzy logic system

The inference engine used in this study is similar to the one described in Vivekanandan et al. (1999). In summary, a probability for each of the four classes (PPT, AP/GC, CHAFF, and CLEARAIR) is computed at every grid point based on values of $\rho_{HV}(0)$, $text(\Phi_{DP})$, $text(Z_{DR})$, and $text(\rho_{HV}(0))$. The probability of a given class (e.g., PPT) for a given variable (e.g., $\rho_{HV}(0)$) is simply “looked up” using the predefined membership functions. The final aggregated probability is computed by adding each of the probabilities found for each variable, where all the variables mentioned above receive equal weight. Finally, the class at a given grid point is determined by finding the maximum of the aggregated probabilities that were computed for each class.

3. RESULTS

Evaluating the performance of a precipitation versus non-precipitation algorithm can be a daunting task. In situ data sets are notoriously limited in their spatial and temporal scales of coverage. Such data sets were not available for this study. An alternative approach is to discriminate hydrometeors from non-hydrometeors for a large, independent data set and then aggregate them over a long period of time. Derived frequency maps reveal spatial patterns of the aggregated classes, which can then be analyzed by a radar expert to determine if the classes were correctly identified.

A day is chosen for analysis (23 June 2005) in which several non-hydrometeor scatterers were noted near the radar. Later in the period, convective echoes developed and moved directly across the radar. The evaluation spans a 24-hour period for which 96 individual scans are available. The classes are derived for each scan using the fuzzy logic system, and their frequency of occurrence is computed over the event duration. All non-hydrometeor classes are grouped together to form the NON-PPT class.

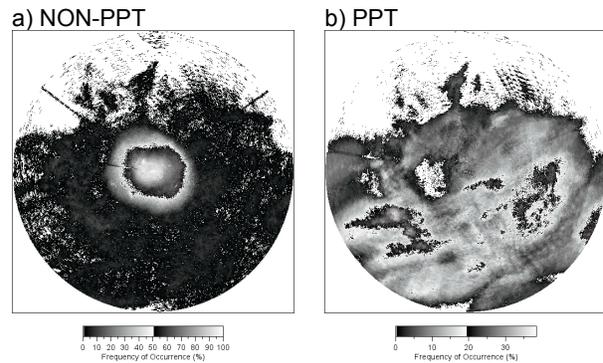


FIG. 3. Frequency of occurrence of a) non-hydrometeors and b) hydrometeors over a 24-hour period.

Fig. 3 shows a remarkable difference between the aggregated class assignments to NON-PPT and PPT. The NON-PPT frequency map has a very large maximum centered on the radar, a region where clear air echoes and ground clutter are common. Also, rays emanating from the radar are noted to the northwest and northeast of the radar. These artifacts are associated with sunset and sunrise. The PPT frequency map, on the other hand, reveals a pattern that is spatially consistent with accumulated rainfall maps.

It is noted in Fig. 3a that there are very low but nonzero frequencies of detecting NON-PPT echoes in regimes dominated by high PPT frequencies. Individual images were examined in order to better understand this possible misclassification of precipitating echoes. It was discovered that the outermost borders of precipitating echoes were occasionally classified as being CLEARAIR. The low values of $\rho_{HV}(0)$ in these regions suggests that signals were not necessarily from CLEARAIR but were more likely from non-precipitating clouds. In any case, the assignment of these occurrences into the NON-PPT class has been considered to be appropriate.

It might appear as though a frequency minimum in Fig. 3b exists at the same location as the maximum in Fig. 3a. If this were the case, then it could be concluded that the algorithm has a less than desirable probability of detecting precipitation. However, closer examination reveals that the precipitation frequency minimum is not located at the center of the image where NON-PPT echoes are likely, but is offset to the west. This minimum is believed to be

a real characteristic of the spatial pattern of precipitation.

It is also interesting to investigate the algorithm's capability of discriminating between different non-hydrometeor types. While the outcome of these results is not as critical as the hydrometeor versus non-hydrometeor discrimination, the identification of biological targets, military and commercial aircraft, and fixed targets may be beneficial to additional agencies and scientific communities. Moreover, these results will provide an expectation on how the fuzzy logic system will perform in hydrometeor species discrimination (e.g., rain versus hail).

A brief analysis of the ability to segregate between AP/GC and CLEARAIR echoes is undertaken. Fig. 4a shows that the maxima in AP/GC frequencies are located very near the radar, with some features resembling anomalous propagation to the west. Close inspection of the image reveals fixed structures that appear routinely on Z_H images. Echoes from CLEARAIR contribute the most to the NONPPT class (Fig. 4b). The spatial pattern of the frequencies near the radar is consistent with expectations. As mentioned previously, cloud echoes have been grouped in the CLEARAIR class, which is not entirely appropriate. Additional classes will be needed in order to isolate these non-hydrometeor species more effectively. It was also noted that several echoes from ground clutter were often mistakenly placed in the CHAFF class. The membership functions of the texture parameters overlap significantly between these two classes (see Fig. 2). There are other parameters that are more capable of distinguishing between CHAFF and AP/GC. The weights on the membership functions would need to be adjusted, as they were optimized for hydrometeor versus non-hydrometeor discrimination in this study. Improved discrimination among non-hydrometeor species is possible through more specific weighting of the membership functions.

4. SUMMARY AND CONCLUSIONS

A scheme has been developed to automatically discriminate between hydrometeors and non-hydrometeors. The unique aspects of this study are the empirical derivation of the membership functions and the evaluation methodology. Typically, the shapes of the membership functions and the values they cover are chosen manually based on simulations, disdrometer measurements, in situ observations, and radar measurements. The method proposed herein is objective and is based on the radar measurements alone. This permits the membership functions to be adaptive and thus appropriate to the radar wavelength and region for which the algorithm will be functioning.

A radar expert first identifies the species of interest. This region is then isolated, and distributions are created for the polarimetric variables using Gaussian kernel density estimation. These functions are then used directly in a fuzzy logic system as membership functions. A judicious selection of variable weights is required by a radar expert.

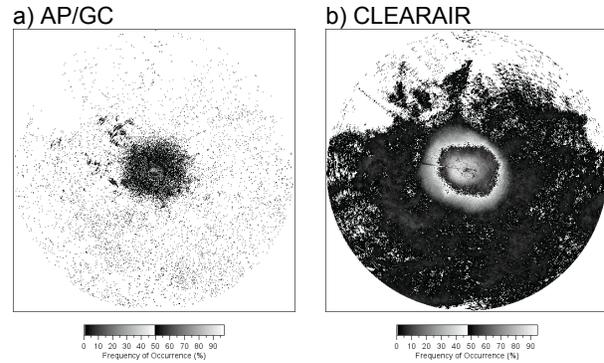


FIG. 4. Frequency of occurrence of echoes from a) anomalous propagation and ground clutter and b) clear air over a 24-hour period.

The proposed scheme was tested on a 24-hour case in which common contaminants were observed near the radar at the beginning of the event. Later, convection developed and moved through the area. Frequency maps were derived for the occurrence of precipitation and other non-hydrometeors. This evaluation methodology is unique in that the spatial patterns of the class frequencies are used to evaluate the performance of the algorithm. This approach covers an entire radar umbrella over a day without an extensive need of in situ verification data sets. Results indicate that the method has a high probability of detecting precipitation with very few false alarms. The capability of the algorithm to distinguish among non-hydrometeor classes was also examined. The performance was acceptable yet suboptimal due to the tuning of the weighting functions for discrimination between hydrometeors and non-hydrometeors. Future work will investigate adapting the weighting of membership functions, deriving K_{DP} membership functions, and utilizing multidimensional functions to accommodate the identification of different particle types (e.g., rain versus hail) and non-hydrometeors (e.g., chaff versus ground clutter).

5. ACKNOWLEDGEMENTS

This work was done in the frame of the PANTHERE Project (Programme Aramis Nouvelles Technologies en Hydrometeorologie Extension et RENouvellement supported by Meteo-France, the French "Ministere de L'Ecologie et du Developpement Durable", the "European Regional Development Fund (ERDF)" of the European Union and CEMAGREF.

6. REFERENCES

- Gourley, J.J., P. Tabary, and J. Parent du Chatelet, 2005a: Data quality of the Meteo-France C-band polarimetric radar. Submitted to *J. Atmos. Oceanic Technol.*
- , 2005b: A new polarimetric method for correcting the effects of attenuation at C-band. This issue.

- Liu, H. and V. Chandrasekar, 2000: Classification of hydrometeors based on polarimetric radar measurements: Development of fuzzy logic and neuro-fuzzy systems, and in situ verification. *J. Atmos. and Oceanic Technol.*, **17**, 140-164.
- Silverman, B.W., 1986: *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 175 pp.
- Straka, J.M., D.S. Zrníc, and A.V. Ryzhkov, 2000: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. *J. Appl. Meteor.*, **39**, 1341-1372.
- Vivekanandan, J., D.S. Zrníc, S.M. Ellis, R. Oye, A.V. Ryzhkov, and J. Straka, 1999: Cloud microphysics retrieval using S-band dual-polarization radar measurements. *Bull. Amer. Meteor. Soc.*, **80**, 381-388.
- Zrníc, D.S., V.M. Melnikov, and A.V. Ryzhkov, 2005: Polarimetric properties of ground clutter. Submitted to *J. Atmos. Oceanic Technol.*