P11B.4

OPTIMIZATION OF THE MATRIX OF WEIGHTS IN THE POLARIMETRIC ALGORITHM FOR CLASSIFICATION OF RADAR ECHOES

HyangSuk Park(1), A. Ryzhkov(2), D. Zrnic(3), Kyung-Eak Kim(1)

(1) Kyungpook National University, Daegu, Korea, (2) CIMMS/OU, Norman, USA, (3) NSSL, Norman, USA

1. INTRODUCTION

Polarimetric radar is uniquely suited for discriminating between different classes of meteorological and nonmeteorological echo. Currently used classification algorithms are based on the principles of fuzzy logic and utilize multiparameter radar measurements with certain weight assigned to each radar variable to account for its classification efficiency (Zrnic and Ryzhkov 1999; Vivekanandan et al. 1999; Liu and Chandrasekar 2000; Zrnic et al 2001; Keenan 2003; Lim et al. 2005). However, no objective justification has ever been given for the choice of these weights in the polarimetric classification scheme.

In this study, we introduce the matrix of weights which characterizes classification power of each variable with respect to every class of radar echo and suggest a methodology for optimizing such a matrix.

The data from several storms observed with the polarimetric prototype of the WSR-88D radar in Oklahoma are used to optimize the matrix of weights in the classification algorithm which is planned for implementation on polarimetric NEXRAD.

2. NEXRAD CLASSIFICATION ALGORITHM

The classification algorithm distinguishes between 10 classes of radar echo: (1) ground clutter / anomalous propagation (GC/AP), (2) biological scatterers (BS), (3) dry aggregated snow (DS), (4) wet snow (WS), (5) crystals of different orientation (CR), (6) graupel (GR), (7) “big drops” (BD), (8) light and moderate rain (RA), (9) heavy rain (HR), and (10) rain / hail mixture (RH). The algorithm utilizes six radar variables: radar reflectivity at horizontal polarization Z, differential reflectivity Z_{DR}, cross-correlation coefficient ρ_{hv}, specific differential phase K_{DP}, and the texture parameters of radar reflectivity SD(Z) and differential phase SD(Φ_{DP}).

The aggregation values or scores for each of 10 classes are determined as

$$A_i = \frac{\sum_{j=1}^{6} W_{ij}\, Q_j\, MF^{(i)}(V_j)}{\sum_{j=1}^{6} W_{ij}\, Q_j} \quad , \qquad (1)$$

where $MF^{(i)}(V_j)$ is a membership function of the j^{th} variable for i^{th} class, W_{ij} is a weight between 0 and 1 assigned to the i^{th} class and j^{th} variable, and Q_j is an element of the confidence vector characterizing instrumental quality of the measurement of the j^{th} variable. The type of radar echo is identified by the maximal aggregation value.

Matrix of weights **W** characterizes classification efficiency of each variable with respect to a particular class. For each pixel or gate, every radar variable is supplemented with its confidence factor Q_j depending on its vulnerability to (a) attenuation, (b) effects of nonuniform beam filling, (c) magnitude of ρ_{hv} (which determines statistical measurement errors of all polarimetric variables), (d) signal-to-noise ratio SNR.

More details on the definition of vector Q, the use of information about the melting layer height, accounting for radar beam broadening, etc., can be found in Ryzhkov et al. (2007).

3. DETERMINATION OF THE MATRIX OF WEIGHTS. THEORY

It is intuitively obvious that the radar variable V_j with heavily overlapped membership functions $MF^{(i)}(V_j)$ for different classes should be given lower weight in the fuzzy logic classification algorithm. Different overlap measures can be used to quantify the degree of overlapping. One of them was suggested by Cho et al. (2006) for discrimination between weather echoes and ground clutter /AP using single-polarization radar. In their study, Cho et al. (2006) utilize the classification rule

$$A_i = \frac{\sum W_j\, MF^{(i)}(V_j)}{\sum W_j} \quad i = 1,2 \quad , \qquad (2)$$

where the elements of the vector of weights **W** are determined as

$$W_j = \frac{1}{A_j \sum A_j^{-1}} \qquad (3)$$

In (3), A_j is the overlapping area between normalized probability density functions $P^{(1)}(V_j)$ and $P^{(2)}(V_j)$ characterizing distributions of the variable V_j for two classes (Fig. 1).

In the case of M classes, we suggest determining the elements of the matrix of weights as follows:

$$W_{ij} = \frac{1}{M-1} \sum_{i \neq k} E_{ik}^{(j)} \quad . \qquad (4)$$

In (4), $E_{ik}^{(j)}$ is a classification efficiency of the variable V_j with respect to classes i and k. The latter is defined as

$$E_{ik}^{(j)} = 1 - OM_{ik}^{(j)} \quad , \qquad (5)$$

Corresponding author address: Alexander Ryzhkov, CIMMS/NSSL, Norman, Oklahoma, USA; email: Alexander.Ryzhkov@noaa.gov

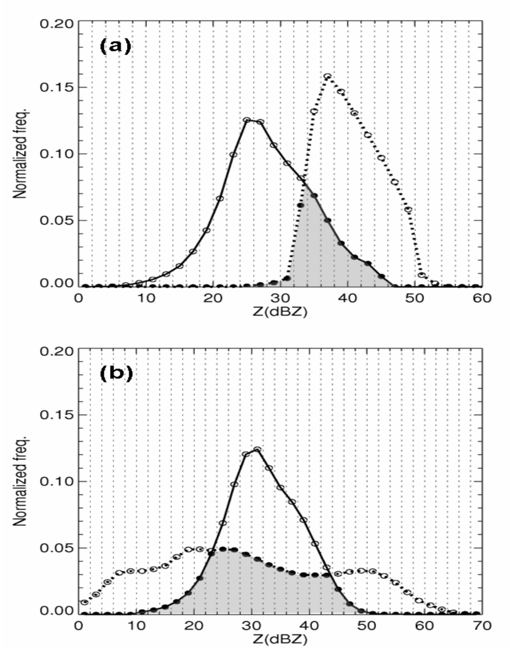

Fig. 1. Illustration of different definitions of the overlap measure

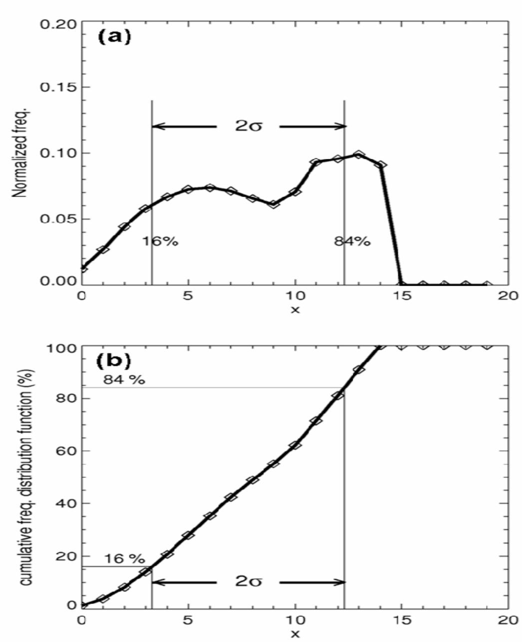

Fig. 2. Definition of the parameter σ

where $OM_{ik}^{(j)}$ is the overlap measure related to classes i and k and variable V_j. In this study, we use three different definitions of the overlap measure.

The first definition is similar to the one utilized by Cho et al. (2006):

$$OM_{ik}^{(j)} = A_{ik}^{(j)} \quad , \qquad (6)$$

where $A_{ik}^{(j)}$ is the overlapping area between normalized probability density functions (PDF) $P^{(i)}(V_j)$ and $P^{(k)}(V_j)$.

According to the second definition, the intersection and union of the functions $P^{(i)}(V_j)$ and $P^{(k)}(V_j)$ are determined and the overlap measure is calculated as the ratio of the corresponding integrals (see e.g., Crum et al. 2006):

$$OM_{ik}^{(j)} = \frac{\int P^{(i)}(V_j) \cap P^{(k)}(V_j)\, dV_j}{\int P^{(i)}(V_j) \cup P^{(k)}(V_j)\, dV_j} \qquad (7)$$

Finally, we introduce the third overlap measure defined as

$$OM_{ik}^{(j)} = \sqrt{2\pi}(\sigma_i \sigma_k)^{1/2} \int P^{(i)}(V_j) P^{(k)}(V_j)\, dV_j \qquad (8)$$

The parameter $\sigma_{i,k}$ means the standard deviation of the probability density function $P^{(I,k)}(V_j)$ if the latter is normal. In a general case of PDF, $\sigma_{i,k}$ is determined as a half of the difference between the 84% and 16% percentiles of the cumulative distribution

$$F^{(i,k)}(X_j) = \int_{-\infty}^{X_j} P^{(i,k)}(V_j)\, dV_j \qquad (9)$$

(see Fig. 2). If the PDFs $P^{(i)}(V_j)$ and $P^{(k)}(V_j)$ are totally identical (100% overlap), then according to all three definitions, $OM_{ik}^{(j)} = 1$ and $E_{ik}^{(j)} = 0$. If the PDFs do not overlap at all, then $OM_{ik}^{(j)} = 0$ and $E_{ik}^{(j)} = 1$.

Analytical expressions for the classification efficiencies $E_{ik}^{(j)}$ and the elements of **W** can be obtained assuming Gaussian shapes of $P^{(i)}(V_j)$ and $P^{(k)}(V_j)$ and using Eq (4), (5), and (8):

$$W_{ij} = \frac{1}{M-1} \sum_{i \neq k} 1 - \left(\frac{\sigma_i \sigma_k}{\sigma_i^2 + \sigma_k^2} \right)^{1/2} \exp\left[- \frac{(x_i - x_k)^2}{2(\sigma_i^2 + \sigma_k^2)^2} \right] \quad (10)$$

where x_i and x_k are modal values of the two PDFs.
Next we illustrate determination of weights given by Eq (10) in three simple cases. In each case, we assume that the modal values x_i are equal for all classes, i.e., the PDFs overlap substantially.

Case 1. Three classes; $\sigma_1 = \sigma$, $\sigma_2 = 2\sigma$, $\sigma_3 = 3\sigma$.
According to (10), $W_{1j} = 0.65$, $W_{2j} = 0.57$, $W_{3j} = 0.62$, i.e., the weights of the variable V_j are higher with respect to the classes with the most narrow and most broad PDF.

Case 2. Three classes; $\sigma_1 = \sigma_2 = \sigma$, $\sigma_3 = 2\sigma$.
In this case, $W_{1j} = W_{2j} = 0.3$, $W_{3j} = 0.6$. This means that if the PDFs of two classes are identical, then the corresponding weights become lower.

Case 3. Four classes; $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$, $\sigma_4 = 2\sigma$.
Adding one more class with the same PDF causes further drop in the corresponding weights: $W_{1j} = W_{2j} = W_{3j} = 0.2$, whereas the weight for the class with distinct (broader) PDF remains intact: $W_{4j} = 0.6$.

4. DETERMINATION OF THE MATRIX OF WEIGHTS. OBSERVATIONS

"True" PDFs for each class and variable should be known in order to construct the matrix of weights **W** according to Eq (4) – (8). The PDFs for 10 classes and 6 radar variables were estimated using the data collected with the polarimetric prototype of the S-band WSR-88D radar during 6 storms observed in Oklahoma. The dataset includes 29 hours of observations.

Radar echo classification has been performed using the latest and most advanced version of the algorithm (Ryzhkov et al. 2007). The algorithm was run with an *apriori* matrix of weights which was constructed based on euristic considerations (i.e., to the best of our knowledge and intuition). The PPI and RHI fields of the radar variables and results of classification were carefully examined, so that only high-confidence class designations were utilized in making “true” PDFs for each class and variable.

One-dimensional normalized frequency histograms which approximate PDFs for 10 classes and 6 variables are shown in Fig. 3. The matrix of weights was determined using Eq (4) for three different definitions of the overlap measure. The matrix **W** corresponding to the first definition of the overlap measure specified by Eq (6) is presented in Table 1.

Table 1. Matrix of weights.

	Z	Z_{DR}	ρ_{hv}	K_{DP}	SD(Z)	SD(Φ_{DP})
GC	0.59	0.56	0.85	0.42	0.68	0.85
BS	0.68	0.56	0.90	0.43	0.38	0.84
DS	0.66	0.62	0.52	0.38	0.20	0.44
WS	0.65	0.53	0.76	0.40	0.19	0.68
CR	0.73	0.54	0.47	0.38	0.23	0.40
GR	0.82	0.49	0.47	0.36	0.24	0.45
BD	0.61	0.49	0.55	0.39	0.23	0.41
RA	0.59	0.45	0.50	0.36	0.19	0.42
HR	0.87	0.75	0.56	0.75	0.20	0.41
RH	0.91	0.47	0.52	0.62	0.22	0.39

It is not surprising that higher matrix elements are associated with the most distinctive PDFs: Z and K_{DP} have highest weights for rain / hail mixture (RH) and heavy rain (HR), Z_{DR} is most informative for heavy rain (HR), whereas ρ_{hv}, SD(Z), and SD(Φ_{DP}) exhibit highest classification capability for nonmeteorlogical scatterers (GC and BS).

Overall classification power of a given variable can be estimated by finding the mean value in the corresponding column of the matrix of weights. According to this criterion, the 6 radar variables are ranked in a following order: Z (0.71), ρ_{hv} (0.61), Z_{DR} (0.55), SD(Φ_{DP}) (0.53), K_{DP} (0.45), and SD(Z) (0.28).

The matrices **W** estimated according to the other two definitions of the overlap measure exhibit the structures which are very similar to the one in Table 1 but with slightly higher absolute values of the matrix elements (not shown). In other words, relative ratios of different elements of **W** are not very sensitive to the choice of the definition of the overlap measure.

REFERENCES

Cho, Y., G. Lee, K. Kim, I. Zawadzki, 2006: Identification and removal of ground echoes and anomalous propagation using the characteristics of radar echoes. *J. Atmos. Oceanic Technnol.,* **23**, 1206 – 1222.

Crum, W., O. Camara, and D. Hill, 2006: Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Trans. Med. Imag.*, **25**, 1451 – 1461.

Keenan, T.D., 2003: Hydrometeor classification with a C-band polarimetric radar. *Aust. Meteor. Mag.*, **52**, 23 – 31.

Lim, S., V. Chandrasekar, and V.N. Bringi, 2005: Hydrometeor classification system using dual-polarization radar measurements: model improvements and in situ verification. *IEEE Trans. Geosci. Rem. Sens*., **43**, 792 – 801**.**

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. Oceanic Technol.,* **17**, 140–164.

Ryzhkov, A., D. Zrnic, P. Zhang, J. Krause, H. Park, D. Hudak, J. Young, J. L. Alford, M. Knight, J. W. Conway, 2007: Comparison of polarimetric algorithms for hydrometeor classification at S and C bands. Analysis of the performance in different climate regions. Preprints, *33rd Conf. on Radar Meteorology.*, CD-ROM, 10.3.

Vivekanandan, J., D. Zrnic, S. Ellis, D. Oye, A. Ryzhkov, J. Straka, 1999: Cloud microphysics retrieval using S-band dual-polarization radar measurements. *Bull. Amer. Meteor. Soc.,* **80**, 381-388.

Zrnic, D., and A. Ryzhkov, 1999 Polarimetry for weather surveillance radars. *Bull. Amer. Meteor. Soc.,* **80**, 389 - 406.

Zrnic, D.S., A.V. Ryzhkov, J. Straka, Y. Liu, and J. Vivekanandan, 2001: Testing a procedure for the automatic classification of hydrometeor types. *J. Atmos. Oceanic Technol*., **18**, 892 – 913.

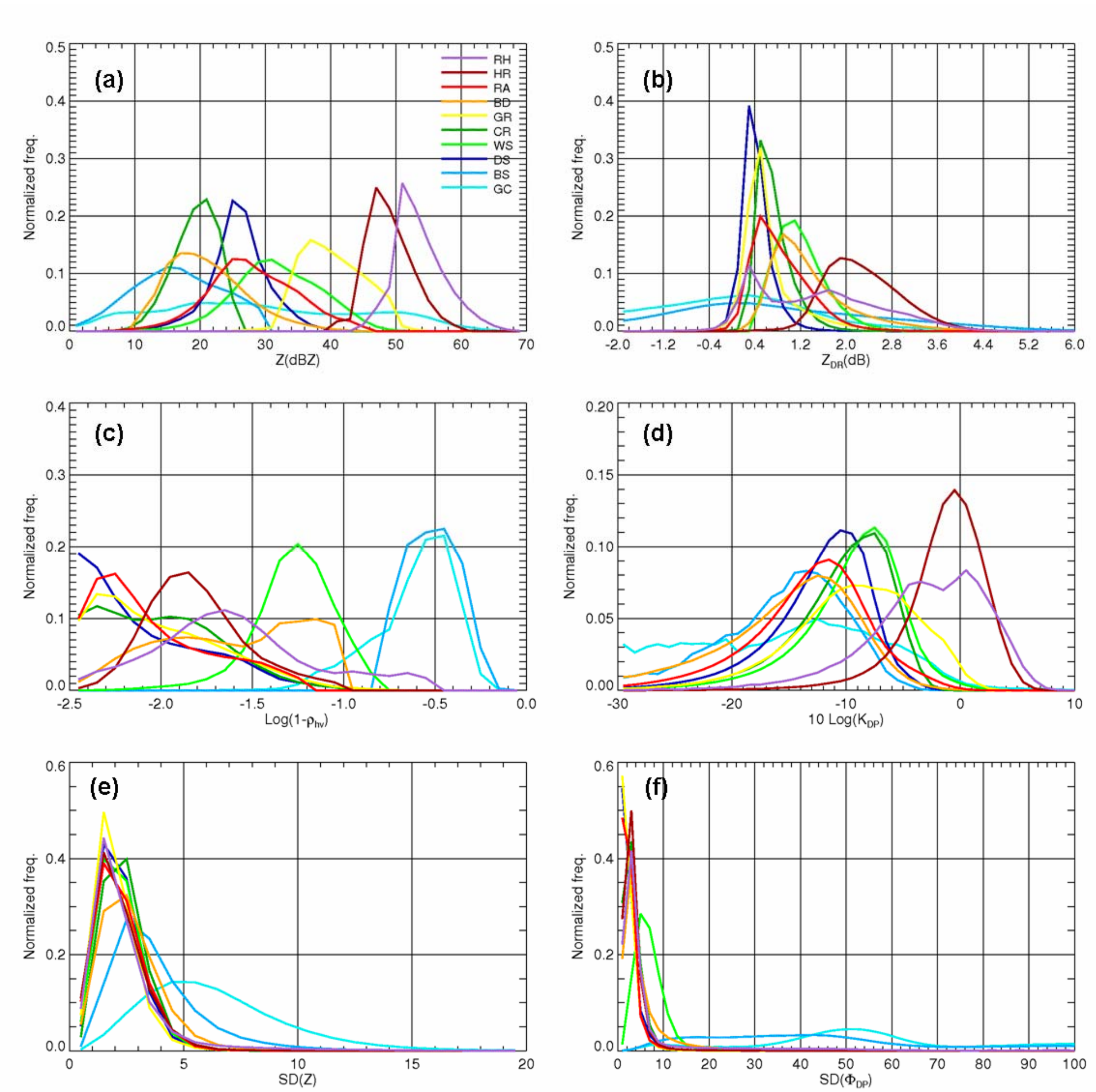

Fig.3 Normalized frequency distribution of radar reflectivity factor Z, differential reflectivity Z_{DR}, cross-correlation coefficient ρ_{hv}, specific differential phase K_{DP}, and texture parameters SD(Z) and SD(Φ_{DP}) for 10 classes of radar echo. GC stands for ground clutter/AP, BS – for biological scatterers, DS - for dry snow, WS – for wet snow, CR – for crystals, BD – for "big drops", RA – for light and moderate rain, HR – for heavy rain, RH – for rain / hail. The normalized histograms are obtained from the data collected by the S-band polarimetric radar in central Oklahoma during 6 storms (29 hours of observations).