1. Introduction

The radar echo classifier (REC) is a data fusion system that uses “fuzzy-logic” techniques (Kosko, 1992) to classify the type of scatterer measured by Doppler radar systems. This algorithm is being installed within the National Weather Service (NWS) Weather Surveillance Radar-Doppler 1988 (WSR-88D) in the Open Radar Product Generator (ORPG) (Saffle and Johnson 1998; Saffle et al. 2001a; Saffle et al. 2001b) as part of the Anomalous Propagation (AP) Clutter Mitigation Scheme (Keeler et al. 1999; Kessinger et al. 2001 and 2002; Ellis et al. 2003). The purpose of the AP Clutter Mitigation Scheme is to improve the WSR-88D radar-derived rainfall estimates by enhancing the quality of the data through the identification and removal of certain contaminants, specifically AP ground clutter. Removal of AP ground clutter is needed since it causes erroneous radar-derived rainfall estimates within the WSR-88D precipitation processing subsystem (PPS; Fulton et al. 1998; O’Bannon 1998) as well as errors in other algorithms and errors in interpretation. The REC determines where AP ground clutter is contaminating the radar base data and where precipitation is occurring to facilitate the automatic detection and removal of AP clutter.

Ground clutter contamination within radar data has two sources: normally-propagated (NP) ground clutter from stationary targets such as buildings, trees, or terrain, and AP ground clutter that arises from particular atmospheric conditions within the planetary boundary layer (i.e., temperature inversions) that duct the radar beam to the ground. Because NP targets are stationary and “unchanging,” they are relatively easy to remove with the use of clutter background (or bypass) maps that reside within the WSR-88D ORPG Human Control Interface (HCI) that shows the Clutter Filter Control (CFC) panel. However, AP clutter can evolve and dissipate as atmospheric conditions change. Presently, recognition of the presence of AP clutter is the responsibility of the NWS forecasters who must manually modify the clutter filter maps to specify additional regions of clutter filter application.

Automatic detection and removal of the AP ground echoes will improve the quality of the base data fields and derived products and relieve forecasters of this tedious responsibility.

Using the three moments of radar base data (reflectivity, radial velocity and spectrum width) as input, various detection algorithms are formulated to make a classification. Currently, three REC algorithms have been designed and tested: the AP detection algorithm (APDA) detects regions of AP ground clutter return (Cornelius et al. 1995; Pratte et al. 1997), the precipitation detection algorithm (PDA) defines convective and stratiform precipitation regions, and the insect clear air detection algorithm (ICADA) defines return from insects in the boundary layer. A sea clutter detection algorithm (SCDA) is under development.

2. Data Sets

Two main sources of radar base data have been used to develop the radar echo classifier: sixty scans from nine WSR-88D systems and many scans from the NCAR S-Pol radar (Keeler et al. 2000) as located in various field projects in the United States and abroad. Kessinger et al (1999a) describes the WSR-88D data set used. Kessinger et al. (1999b, 2001, 2002), Kessinger and Van Andel (2001), and Ellis et al. (2003) describe the S-Pol data sets.

The WSR-88D is a 10 cm wavelength, single-polarization radar. The S-Pol radar is a 10 cm wavelength, dual-polarization radar. S-Pol uses a four-pole elliptical, high pass ground clutter filter with passband edges of \( \pm 0.5 \text{ m s}^{-1} \). The WSR-88D clutter filter for low suppression has passband edges of \( \pm 1.2 \text{ m s}^{-1} \); for medium suppression, passband edges are \( \pm 1.6 \text{ m s}^{-1} \). Because of its comparatively narrow width, the S-Pol ground clutter filter removes less data from the power spectra when compared to a WSR-88D.

The various locations where the S-Pol radar has been deployed include Brazil in January and February 1999, eastern Colorado in June and July 2000, the Pacific coast of Washington state in January and February 2001, the Cascade Mountains of Oregon in December 2001, and the Oklahoma panhandle in May and June 2002. Beginning in the summer of 2000, the REC has been run operationally on the S-Pol radar during field operations to facilitate algorithm optimization and to
ensure that the algorithms are computationally efficient.

The switch from using WSR-88D radar data to using the S-Pol radar data to develop and optimize the REC algorithms occurred because of the development of the particle identification algorithm (PID; Vivekanandan et al. 1999), described below.

3. Methodology

To develop the algorithms for the REC, several steps must be done to prepare the data. The first step is to “truth” the data such that the characteristics of each scatterer type can be examined to find those that discriminate one scatterer from another. The characteristics of each scatterer are examined through the “feature fields”. The feature fields are quantities (e.g., mean, median, standard deviation) that are derived from the radar base data fields of reflectivity, radial velocity and spectrum width. Next, histograms of the feature fields are made to examine the distribution of values between scatterer types. Finding those features that discriminate one scatterer type from another is a key task. Once the feature fields are selected for each scatterer type, “membership functions” can be devised, based upon the distribution of values. Membership functions will be more fully described below.

1.1 Truth Fields

“Truthing” is an important activity. Comparison with truth as defined by expert humans or by the particle identification (PID) algorithm is crucial in optimizing the performance of the fuzzy logic algorithm.

For the WSR-88D data set, the expert defines the truth field by examining the various radar fields and their derived products using the NCAR SOLO program (Oye et al., 1995). Only the 0.5 deg elevation angle is truthed but the other elevation angles are examined for additional information. All regions of the 0.5-degree scan are characterized as ground clutter, precipitation, or insect clear air return. Figure 1 shows an example of the expert truthed field.

Typically, AP clutter is characterized by a near-zero radial velocity, low spectrum width, high texture of the reflectivity field, as well as other characteristics. Further, the expert can examine the scan at the next two elevation angles (1.5 and 2.3 degrees) to determine if the radar echo in question has vertical continuity. A lack of vertical continuity does not guarantee that the echo is AP clutter, but gives more information into the decision-making process. Once the expert has determined where the AP clutter is, SOLO is used to draw a polygon around the region and the appropriate value is set within the truth field.
Figure 2. An example of how the multiple categories of the a) particle identification algorithm (PID) output are converted to b) an objective truth field. Table 1 lists the category numbers for each color value in a). See Figure 1 for the color key for b). NCAR S-Pol data on the 0.5-degree elevation angle scan are shown on 23 June 2000 at 0105 UTC from the STEPS field experiment.

<table>
<thead>
<tr>
<th>Echo Type</th>
<th>Number</th>
<th>REC Algorithm Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud</td>
<td>1</td>
<td>Not used</td>
</tr>
<tr>
<td>Drizzle</td>
<td>2</td>
<td>Not used</td>
</tr>
<tr>
<td>Light rain</td>
<td>3</td>
<td>Precipitation Detection Algorithm</td>
</tr>
<tr>
<td>Moderate rain</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Heavy rain</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Hail</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Rain and hail mixture</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Graupel and small hail</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Graupel and rain</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Dry snow</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Wet snow</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Ice crystals</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Irregular ice crystals</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Super-cooled liquid droplets</td>
<td>14</td>
<td>Not used</td>
</tr>
<tr>
<td>Flying insects</td>
<td>15</td>
<td>Insect Clear Air Detection Algorithm</td>
</tr>
<tr>
<td>Flying birds</td>
<td>none</td>
<td>Not used</td>
</tr>
<tr>
<td>Ground clutter</td>
<td>17</td>
<td>AP Detection Algorithm</td>
</tr>
<tr>
<td>Second trip echo</td>
<td>16</td>
<td>Removes echo</td>
</tr>
</tbody>
</table>

Table 1. The radar scatterers identified by the particle identification (PID) algorithm are listed with the corresponding radar echo classifier (REC) algorithm that uses the PID output as the objective truth field.

Of course, a basic limitation of defining the truth in this manner is the lack of independence between the truth field and the radar fields used as input into the REC. A preferable methodology is to have an independent source for determination of truth regions. The data available from the S-Pol dual-polarimetric radar provides an opportunity to obtain an independent determination of the truth field through use of the particle identification (PID) fuzzy logic algorithm (Vivekanandan et al. 1999). The PID classifies various echo types, including ground clutter, precipitation type and insects. A sounding is input into the PID. Echo types that are currently being classified by the PID are listed in Table 1 with the REC algorithm that uses the echo type as the truth field also listed.

To construct the truth field for evaluation of the REC algorithms, PID categories can be grouped. For the APDA, the one category of “ground clutter” defines the truth field. For the PDA, the categories from “light rain” through “irregular ice crystals” are used to define the location of precipitation echoes. For the ICADA, the “flying insects” category is used to define the truth. Figure 2 illustrates how the various PID categories are grouped to form the truth field.

Two disadvantages of using S-Pol data exist. The first is that data are not collected beyond the first trip region, as is done in the WSR-88D systems. The
second is that there are small differences in the radar system characteristics between WSR-88D and the S-Pol. Therefore, all algorithm development done with S-Pol data must be verified with the WSR-88D data set.

1.2 Histograms of Data Characteristics

To determine the characteristics of the various echo types, histogram plots are made for the feature fields using the truth field to distinguish between the types of radar echoes. Figure 3 shows examples of histograms of two variables currently used in the REC as derived from data taken by the Dodge City, KS (KDDC) WSR-88D radar. Fields shown include the texture of the reflectivity (TDBZ) and the mean radial velocity (MVE). These figures show that clutter is characterized by high values of TDBZ and by near-zero values of MVE while precipitation plus clear air return is characterized by low values of TDBZ and that MVE has a broad distribution.

After construction of the histograms, they are examined to find the best discriminators for the echo type of interest. This process is used to define the features used in the REC algorithms. The histograms also aid in defining the membership functions of the feature fields used in the algorithm.

Figure 3. Histogram plots showing the number of range gates containing each feature field (expressed in hundredths of percentage points of the total number of gates) within clutter (left column) and within precipitation plus clear air return from insects (right column) as determined by the truth field. Fields shown are a) texture of the reflectivity (TDBZ; dBZ) within clutter, b) the TDBZ within precipitation and clear air return, c) the mean velocity field (MVE; m s⁻¹) within clutter, and d) the MVE within precipitation and clear air return. Data were taken from one scan of the Dodge City, KS, WSR-88D at 0.5 degree elevation.
4. Radar Echo Classifier

A schematic of the REC is shown in Figure 4. The base data fields of reflectivity (Z or DBZ), radial velocity (V), and spectrum width (W) are input into the “feature generator” for calculation of derived fields that are called “features”. Each of the REC algorithms uses a different combination of feature fields as input. A “membership function” is applied to the values of the feature fields to scale them to match the characteristics of the echo type under consideration. Membership functions are stepwise linear functions that scale the feature fields to have values between zero and unity. Output fields from this process are termed “interest fields”. When an interest field has a value of 1.0, high likelihood exists that the echo type matches the desired characteristics; likewise, an interest value of 0.0 indicates no likelihood that the echo type matches the desired characteristics. Once the interest fields are calculated for all the algorithm feature fields, a weighted mean of all interest outputs is computed. The final step is to apply the appropriate threshold to identify regions that belong in a particular class. A threshold of 0.5 is generally used.

Numeric Python is the computing language used for the REC (Van Andel, 2000 and 2001).

1.3 Feature Field Generation

In the REC, all features are calculated as a first step, and then each detection algorithm selects the pertinent features. With one exception, the features are computed over a small, local region typically defined as two beams on either side of the current beam and ±1 km (Doppler fields) or ±2 km (reflectivity fields) in range from the current gate. The exception is the vertical difference of the reflectivity (VdZDiff) that is computed as a gate-to-gate difference.

The “texture” feature is the mean squared difference of the input field (X) and is calculated as shown in Eq. 1.

\[
\text{Texture} = \left( \frac{1}{N_{\text{beams}}} \sum_{i=2}^{N_{\text{beams}}} \left( X_{i} - X_{i-1} \right)^2 \right) / N
\]

Eq. 1

As described above, the vertical difference of the reflectivity (VdZDiff) field is computed as a gate-to-gate difference of the reflectivity values between the second elevation angle (typically 1.5 degrees) and the first elevation angle (typically 0.5 degrees) as shown in Eq. 2.
\[ V_{dZDiff} = dBZ_{\text{upper}} - dBZ_{\text{lower}} \]  \hspace{1cm} \text{Eq. 2}

Other feature fields include the mean reflectivity (MDZ), median radial velocity (MDVE) and the median spectrum width (MDSW), the standard deviation of the radial velocity (SDVE) and the standard deviation of the spectrum width (SDSW). All are calculated over the local area.

Two additional reflectivity features were added to the REC and are called "SPINchange" and "SIGN". The SPINchange variable (also called SPIN) is taken from Steiner and Smith (2002) and indicates the number of inflection points within the gate-to-gate reflectivity difference field, expressed as a percentage of all possible differences, that exceed the minimum difference (DBZthresh) allowed. However, SPINchange_counts is not incremented unless there has been a change in the sign of the reflectivity gradient (i.e., an inflection point exists). For example, if reflectivity is increasing along the beam and then decreases, SPINchange_counts is only incremented at the gate where the decrease begins. Steiner and Smith (2002) use a value of DBZthresh=2.0 for the WSR-88D data. For the APDA (using S-Pol), the value of DBZthresh=11.0. This higher threshold better discriminates ground clutter from precipitation echo.

The SIGN feature is the mean sign of the reflectivity change and was suggested by Tim O’Bannon (Radar Operations Center; personal communication). The SIGN is computed as the average of the signs of the change in reflectivity from gate to gate within the local area. The sign of reflectivity change for each gate is considered to be +1, 0, or -1 if the difference in reflectivity between that gate and the gate immediately nearer the radar is positive, zero or negative, respectively. A value of SIGN near +1 or -1 can indicate strong reflectivity gradients such as those within convective precipitation. Smaller values (near zero) are expected within AP clutter.

1.4 Membership Functions

Using the feature field histograms (Figure 3), membership functions are devised such that the characteristic that matches the scatterer type has a maximum likelihood value. For instance, ground clutter typically has radial velocity values near zero m s\(^{-1}\). Therefore, the appropriate membership function will scale values of zero m s\(^{-1}\) to a value of unity, i.e., the maximum degree of likelihood. Values of radial velocity far from zero, such as 10 m s\(^{-1}\), will be scaled to have a likelihood value of 0. Defining the appropriate velocity value where zero likelihood should begin on the membership function depends on the data distribution for the feature. For features having a narrow distribution (Figure 3c), the membership function will also have a narrow distribution. Features having a broad distribution (Figure 3a) will have membership functions with a broad distribution. The membership functions for each of the REC algorithms are shown below.

5. The REC Algorithms

In this section, the algorithms that currently comprise the REC are discussed. Table 2 lists the feature fields and the associated weights used for the algorithms.

1.4.1 ANOMALOUS PROPAGATION DETECTION ALGORITHM (APDA)

The feature fields used by the APDA include: the "texture" of the reflectivity field (TDBZ), the median radial velocity field (MDVE), the median spectrum width field (MDSW), the standard deviation of the radial velocity field (SDVE), the SPIN field and the SIGN field (Table 2). Membership functions are shown in Figure 5.

2.4.1 PRECIPITATION DETECTION ALGORITHM (PDA)

The precipitation detection algorithm (PDA) uses the feature fields of the standard deviation of the radial velocity (SDVE) and the spectrum width (SDSW), the texture of the reflectivity field (TDBZ), the mean reflectivity (MDZ) and the SPIN and SIGN reflectivity variables (Table 2). Membership functions are shown in Figure 6. The membership functions shown were used at field projects during winter conditions. Some retuning of the membership functions may be needed for seasonal conditions. In cases of strong clear air return, a reflectivity threshold of 10 dBZ on the final interest field improves performance. Distinguishing precipitation from insect clear air return (and vice versa) can be difficult when clear air return is strong.

<table>
<thead>
<tr>
<th>APDA</th>
<th>PDA</th>
<th>ICADA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
<td>Wt</td>
<td>Feature</td>
</tr>
<tr>
<td>TDBZ</td>
<td>1</td>
<td>TDBZ</td>
</tr>
<tr>
<td>MDSW</td>
<td>1</td>
<td>SDZW</td>
</tr>
<tr>
<td>SDVE</td>
<td>1</td>
<td>SDVE</td>
</tr>
<tr>
<td>MDVE</td>
<td>1</td>
<td>MDZ</td>
</tr>
<tr>
<td>SPIN</td>
<td>1</td>
<td>SPIN</td>
</tr>
<tr>
<td>SIGN</td>
<td>1</td>
<td>SIGN</td>
</tr>
</tbody>
</table>

Table 2. Feature fields and weights used for each interest field of the AP detection algorithm (APDA), the precipitation detection algorithm (PDA) and the insect clear air detection algorithm (ICADA).
Figure 5. Membership functions for the AP clutter detection algorithm (APDA).

Figure 6. Membership functions for the precipitation detection algorithm (PDA) as used during the IMPROVE-I and -II field programs.
3.4.1 INSECT CLEAR AIR ECHO DETECTION ALGORITHM

The insect clear air detection algorithm (ICADA) uses the feature fields of the texture of the radial velocity (TVE), the mean spectrum width (MSW), the texture of the reflectivity field (TDBZ), the mean reflectivity (MDZ), and the SPIN and SIGN reflectivity variables (Table 2). Membership functions are not shown. This algorithm is the least robust of the three and must be tuned for seasonal conditions.

6. Field Testing of the REC

The REC has been field tested on S-Pol during field operations since the summer of 2000. Within the last year, two field programs occurred and the REC was run in real-time for each.

The International H₂O Program (IHOP) field experiment was held in the Oklahoma and Texas panhandles during May and June of 2002. The S-Pol radar was located near Beaver, OK. The terrain has little relief in this region; however, AP conditions frequently occur after the passage of large thunderstorms.

During late 2001, the REC algorithms were tested during the Improvement of Microphysical Parameterization through Observational Verification Experiment II (IMPROVE-II) field program that was conducted by scientists from the University of Washington. The experiment was held from late November-December 2001 in the Oregon Cascade Mountains near Sweet Home, OR. The S-Pol radar was located on the summit of a mountain, surrounded by higher terrain. Ground clutter contamination and beam blockage were considerable at the lowest elevation angles.

An example of the REC performance during the IHOP experiment is shown in Figure 7. This case occurred on 16 June 2002 at 0000 UTC. An extensive area of AP clutter has formed near the radar following the passage of a gust front. Results from the REC are shown with the reflectivity and radial velocity fields. As is seen in Figure 7c, the APDA detects the several regions of AP clutter quite well (denoted by the green shading). The APDA has strong performance when the clutter regions are large and homogeneous. The PDA, shown in Figure 7d, detects the convective storms quite well (denoted by the gold shading). A region of insect clear air return is indicated by the red arrow and is a region of false classification by the PDA. Differentiating between clear air return and precipitation can be difficult when clear air return has reflectivity values in excess of 10 dBZ.

Figure 8 shows one of the cases from IMPROVE-II with output from the APDA and the PDA. The data were collected on 5 December 2001 at 0126 UTC during a large scale, stratiform rain event. As seen in Figure 8c, the APDA detects the ground clutter when it is not embedded within precipitation. When ground clutter and precipitation coexist, the APDA has difficulty detecting the ground clutter because the feature calculations are calculated over a local area. This is a known limitation of the APDA and efforts are underway to improve the detection performance for small regions of clutter. The PDA thresholded output is shown in Figure 8d. This algorithm performs well at detecting regions of precipitation. It falsely detects some regions of ground clutter as precipitation. Efforts are underway to improve the PDA performance in mixed precipitation and clutter conditions.

Ellis et al. (2003) shows other examples of REC performance from the IMPROVE-II and IHOP field programs.

7. Summary and Future Work

A description of the fuzzy logic, data fusion system called the radar echo classifier (REC) has been presented. The REC discriminates between various types of radar return to define AP ground clutter, precipitation, and insect clear air return. The REC is a modular system that allows new detection algorithms to be added in a straightforward fashion.

The APDA has been deployed within the WSR-88D ORPG build 2 in October 2002. This allows the digital output from the APDA to be available to users to assist in the removal of AP ground clutter. Ellis et al. (2003) discusses the ORPG implementation of the REC in detail.

A sea clutter detection algorithm is under development. Results from this new algorithm will be shown at the conference.

Algorithms that estimate the quality of the REC algorithm output (so-called “confidence algorithms”) are under development.

8. Acknowledgements

The NOAA Radar Operations Center (ROC) sponsors this research.

9. References


Figure 7. S-Pol data from the IHOP field experiment on 16 June 2002 at 0000 UTC. Fields shown include: a) reflectivity (dBZ), b) radial velocity (m s\(^{-1}\)) with values near zero shaded cyan, c) thresholded APDA (green) and d) thresholded PDA (gold). The red arrow in d) denotes region of clear air return that is falsely detected as precipitation. The 0.0-degree elevation angle is shown. Range rings are at 30 km intervals.
Figure 8. S-Pol data from the IMPROVE-II field experiment on 5 December 2001 at 0126 UTC. Fields shown include: a) reflectivity (dBZ), b) radial velocity (m s\(^{-1}\)) with values near zero shaded cyan, c) thresholded APDA (green) and d) thresholded PDA (gold). The red boundaries are encircling regions of ground clutter caused by high terrain. The 1.5-degree elevation angle is shown. Range rings are at 30 km intervals.


