Fuzzy Rule–Based Approach for Detection of Bounded Weak-Echo Regions in Radar Images

NIKHIL R. PAL

Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta, India

ACHINTYA K. MANDAL

Gitanjali Net, Visva-Bharati University, Santiniketan, India

SRIMANTA PAL AND JYOTIRMAY DAS

Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta, India

V. LAKSHMANAN

National Severe Storms Laboratory, University of Oklahoma, Norman, Oklahoma

(Manuscript received 4 February 2005, in final form 27 February 2006)

ABSTRACT

A method for the detection of a bounded weak-echo region (BWER) within a storm structure that can help in the prediction of severe weather phenomena is presented. A fuzzy rule–based approach that takes care of the various uncertainties associated with a radar image containing a BWER has been adopted. The proposed technique automatically finds some interpretable (fuzzy) rules for classification of radar data related to BWER. The radar images are preprocessed to find subregions (or segments) that are suspected candidates for BWERs. Each such segment is classified into one of three possible cases: *strong* BWER, *marginal* BWER, or *no* BWER. In this regard, spatial properties of the data are being explored. The method has been tested on a large volume of data that are different from the training set, and the performance is found to be very satisfactory. It is also demonstrated that an interpretation of the linguistic rules extracted by the system described herein can provide important characteristics about the underlying process.

1. Introduction

Supercell thunderstorms are perhaps the most violent of all thunderstorm types and are capable of producing severe winds, large hail, and weak-to-violent tornadoes (Dostalek et al. 2004). In general, however, the supercell class of storms is defined by a persistent rotating updraft (i.e., mesocyclone), which promotes the organization, maintenance, and severity of thunderstorm. The weather surveillance radars used by the National Weather Service (Crum and Alberty 1993) provide vivid pictures of the mesocyclonic activities. The

interaction between the updrafts and the vertically sheared environment strongly controls the degree of the organization of a storm and the severity of convection. Supercells and tornadoes are associated with moderate-to-strong vertical wind shear and moderate-to-high instability.

A bounded weak-echo region (BWER) is a radar signature within a thunderstorm characterized by local minima in the radar reflectivity at low levels, which extends upward into and is surrounded by the higher reflectivities aloft. This feature is associated with a strong updraft and is almost always found in the inflow region of a thunderstorm (Markowski 2002; Cotton and Anthes 1989). In fact, a BWER is a representative of a local storm that develops in a strongly sheared environment and tends to a steady-state circulation.

The weather surveillance radars used by the National

Corresponding author address: Prof. Nikhil R. Pal, Electronics and Communication Sciences Unit, Indian Statistical Institute, 203 B.T. Road, Calcutta 700108, India.
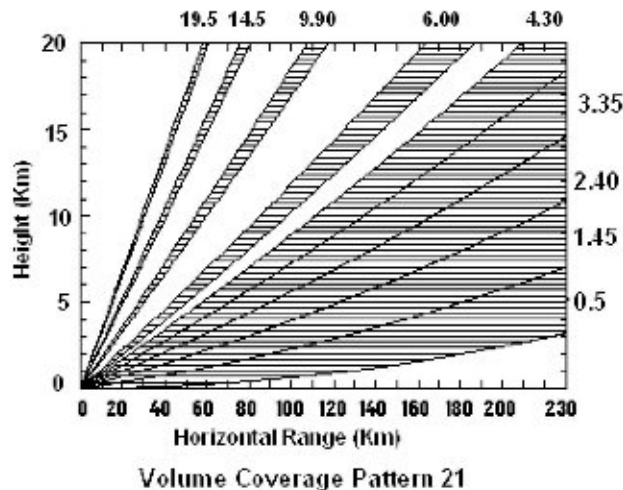E-mail: nikhil@isical.ac.in

FIG. 1. Volume coverage of a WSR-88D radar with a bandwidth of 0.95° and nine elevation scans. (From Smith 1995.)

FIG. 2. Vertical cross section of a BWER. The contours represent constant radar reflectivity (Lakshmanan and Witt 1996; Lemon 1980).

Weather Service (Crum and Alberty 1993) scan through thunderstorms starting at a low elevation angle (0.5°), and after completing a full 360° azimuthal sweep they progressively increase the elevation angle until an upper limit (19.5°) is reached. This is shown in Fig. 1. As the radar scan continues, a BWER first appears as a region of relatively low reflectivities surrounded by higher reflectivities at a lower elevation angle (Krauss and Marwitz 1984). Then, at higher elevation angles it becomes "capped" by a broad region of high reflectivity. The vertical cross section of a BWER is depicted in Fig. 2, and Fig. 3 depicts the horizontal cross section of a strong BWER.

There are several factors that make the radar signature of a BWER rarely appear very clear. As the distance between the radar and the storm increases, the ability of the radar to properly sample small-scale features within the storm, such as a BWER, becomes more difficult. This is so because the radar-sampling volume becomes larger. There is another problem with rapidly moving storms. By the time the radar scans upward through the storm, the higher-altitude capping region of the BWER may move significantly with the storm and may no longer be located over the relative reflectivity minimum detected at a lower altitude (Lakshmanan and Witt 1996). There is also an error associated with the vertical height measured by the radar (Howard and Gourley 1995), and this error varies with weather conditions.

Smalley et al. (1995) have proposed a method based on the 2D structure but did not consider the 3D structure of a BWER in the decision process. The success rate of this method (Smalley et al. 1995) is limited because it detects all local minima in the reflectivity field
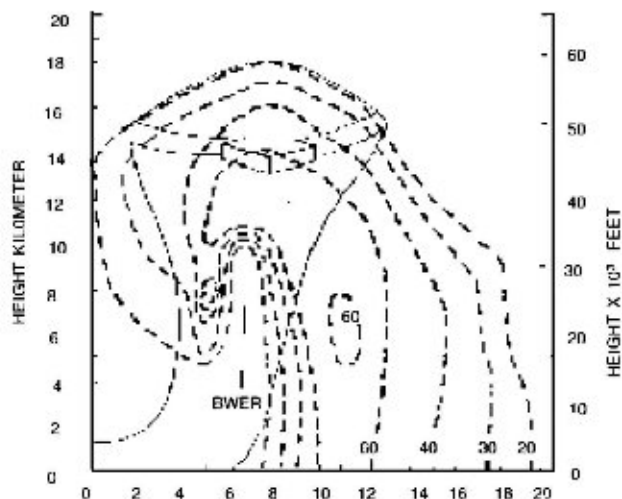
of the radar data. In this 2D scheme, the false-alarm rate is very high.

Fuzzy logic has been used to improve the performance of meteorological information processing systems in the past, notably in gust front (Delanoy and Troxel 1993) and also in BWER (Lakshmanan and Witt 1996) detection. Because of the various uncertainties associated with the appearance of a BWER in radar images, a fuzzy rule–based classification scheme is expected to work well.

An improved method for the detection of BWERs was proposed by Lakshmanan and Witt (1996). In this method they have considered the 3D structure of the
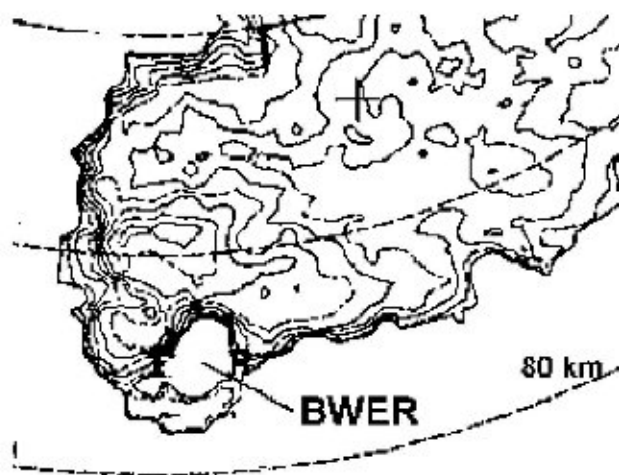

FIG. 3. Horizontal cross section of a strong BWER. The contours represent constant radar reflectivity (Lakshmanan and Witt 1996; Krauss and Marwitz 1984).

BWER in their decision process. Their detection scheme is designed based on the concept of fuzzy logic. Lakshmanan and Witt (1996) computed 23 features on each suspected region obtained from the radar image, and these features are characterized by a fuzzy scale. They also assigned a weight to each feature. The decision is made based on a weighted average of features. Lakshmanan (2000) have used genetic algorithms (GAs) for tuning of the BWER detection algorithm.

In this paper we propose a fuzzy rule–based model for the detection of BWERs. We develop an automatic scheme for the extraction of fuzzy rules for classification of the suspected regions. We have used the dataset that was extracted and used by Lakshmanan and Witt (1996) and Lakshmanan (2000). Each feature vector in the dataset represents a region or segment suspected to contain a BWER or includes a part of a BWER. [We use the words *region, subregion,* and *segment* interchangeably to represent a small portion of a radar image that may contain a BWER. Note that, the word "segment" has a different usage in the Storm Cell Identification and Tracking algorithm for the Weather Surveillance Radar-1988 Doppler (WSR-88D).] We first generate a set of initial fuzzy rules from the dataset using exploratory data analysis techniques to classify each candidate region into one of three classes—*strong, marginal,* and *no* BWER. The initial rule set is then tuned using a gradient search for performance improvement. We achieved an improved performance on the training set as well as on the test dataset relative to the results reported by Lakshmanan and Witt (1996) and Lakshmanan (2000).

We describe our proposed model in section 2. The results are shown in section 3. Section 4 concludes the paper.

2. Proposed model

The fuzzy rule–based classification model, a conceptual model to classify objects, is based on approximate reasoning. The fuzzy set theoretical classification framework provides a degree of support to each potential class. A set of fuzzy rules is used to describe a particular class. The rules are defined on some features, which are computed from the radar images. For any new data (test data), the degree of the match of the features with each fuzzy rule is computed. The class label associated with the rule having the strongest match (i.e., the highest firing strength) defines the class of the new data point.

The proposed method consists of the following three steps: 1) generation of the training and test datasets (discussed in section 2a), 2) generation of an initial

fuzzy rule base (discussed in section 2b), and 3) refining (tuning) of the rule base (discussed in section 2c).

a. Training and test data

Lakshmanan and Witt (1996) and Lakshmanan (2000) have collected data over 5 days, and we use the same data. The radar scans obtained on 4 days (25 May 1996, 21 April 1996, 2 June 1995, and 7 May 1995) contain 186 strong and 59 marginal BWERs. For the fifth day (1 June 1995) there was no BWER. Figure 4 shows a typical WSR-88D image from the Lubbock, Texas, (KLBB) radar at 2136 UTC 25 May 1996. The polar data were projected on a plain with a 1 km × 1 km uniform resolution Cartesian grid. The plane is tangential to the earth's surface at the radar location. The Cartesian grid was limited to 256 km in range. Figure 4 has a BWER as is marked on it. The typical time to process a volume scan is about 30 s on a Sun workstation.

These radar images are preprocessed to find segments (subregions) that are suspected candidates for BWERs (Lakshmanan and Witt 1996). A segment is identified by looking for local minima in the radar elevation scans. Contiguous range gates that belong to local minima are connected to form a candidate segment. These segments are then labeled as having either strong, marginal, or no BWER. Each such segment is represented by a set of features, such as the number of pixels in the segment, the maximum and minimum values of the radar reflectivity, and so on. These data are divided into two parts—one for training of our proposed model and the other for testing the performance of the trained model. The training dataset includes 3 days of data (out of 4 days) with BWERs, along with 50% of the fifth day's data without BWER. The observations of the remaining 1 day with BWERs along with the remaining 50% of data of the fifth day are used for testing. The training dataset contains 1479 strong, 735 marginal, and 4501 no BWER candidate segments, and the test dataset contain 2113 strong, 642 marginal, and 4501 no BWER segments. There are many segments with no BWER, but we have randomly selected 4501 segments for the training set and the same number of segments for the test set. We shall denote the training set as $\mathbf{X}^{\text{Tr}}$ and test set as $\mathbf{X}^{\text{Te}}$.

Each segment is represented by 23 features, which can be divided into two major groups: geometric characteristics (contains 6 features) and BWER characteristics (contains 17 features).

We now provide a description of each feature. The six geometric features are cx, cy, minimum_x, minimum_y, maximum_x, and maximum_y, where (cx, cy) is the coordinate of the center of the BWER. Here
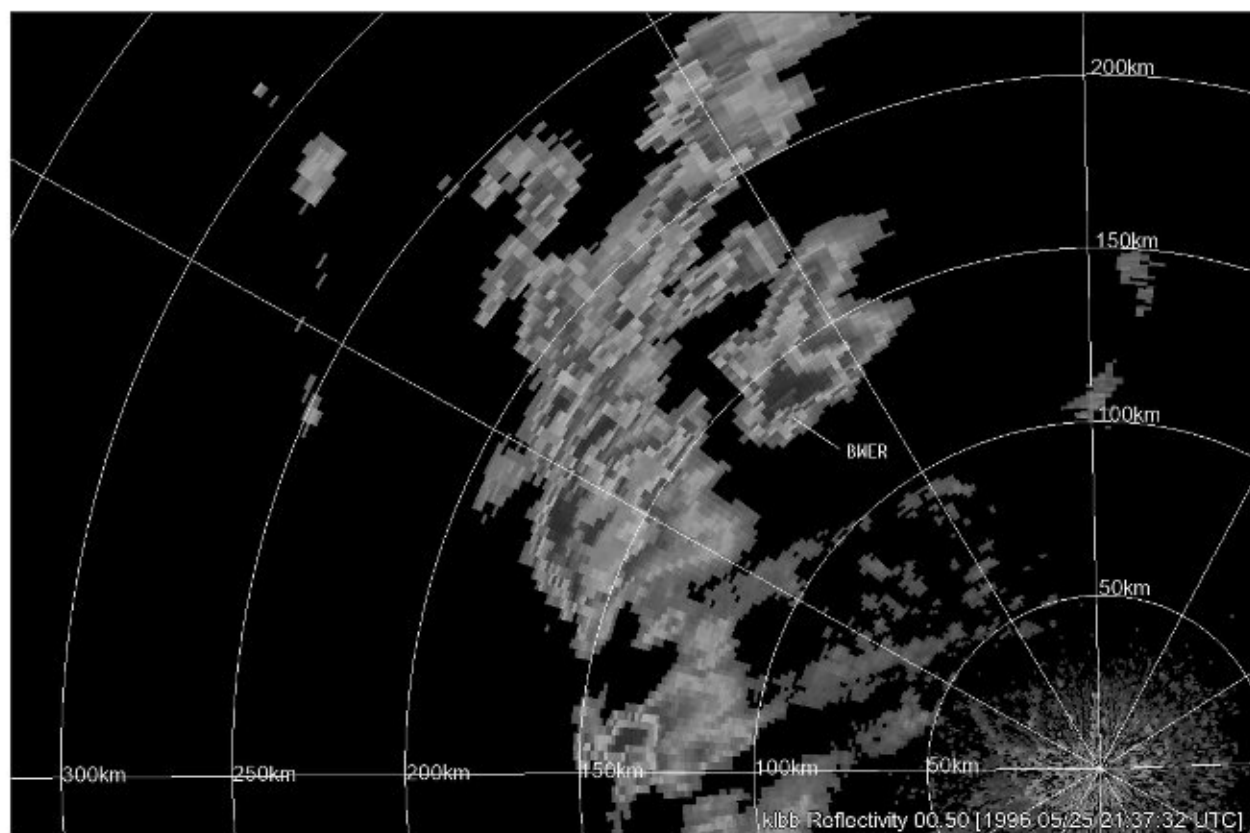
F<small>IG</small>. 4. A typical radar image from KLBB radar.

(minimum_x, minimum_y) and (maximum_x, maximum_y) are the coordinates of the top-left and the bottom-right corners of the smallest rectangle containing the BWER, respectively. The coordinates are specified on a grid that is tangential to the earth's surface at the radar location. The grid is at the resolution of the radar range gates [1 km for the Next Generation Weather Radar (NEXRAD)]. The radar elevation scans are projected to that tangential grid.

The remaining 17 BWER characteristics are number _of_pixels, maximum_rfl, minimum_rfl, average_rfl, maximum_bound, minimum_bound, average_bound, number_of_bounds, maximum_cap, minimum_cap, average_cap, height_of_the_BWER, best_height, low_VIL, prev_conf, cov_vol, and sweep.

Here, number_of_pixels is the number of pixels in the segment of the radar image. The features maximum_rfl, minimum_rfl, and average_rfl represent the maximum, the minimum, and the average reflectivity values in the segment (measured in dBZ).

The maximum_bound, minimum_bound, and average_bound measure some bounding values of the reflectivities that surround the weak-echo region in the radar elevation scan. The maximum_bound is defined

as the highest reflectivity value in the immediate surrounding neighbor of the candidate 2D weak-echo region. Similarly, the minimum_bound (average_bound) is defined as the lowest (average) reflectivity value computed from the immediate neighbors of the candidate 2D weak-echo region. These three features are measured in dBZ. The number_of_bounds is the number of such boundary pixels that surround the weak-echo region in a radar elevation scan.

The cap at a pixel of the weak-echo region is the highest reflectivity value in any of the elevation scans above the weak-echo region. The maximum_cap of a candidate segment is the maximum reflectivity value of the caps of each of its constituent pixels. Similarly, the minimum_cap is the minimum reflectivity of the caps of each of its constituent pixels. The average_cap is the average value of the caps, averaged over all pixels that together form the weak-echo region. These features are also measured in dBZ.

There are three features that capture some 3D properties: height_of_the_BWER, best_height, and low_VIL (vertically integrated liquid). A BWER is a 3D structure that is formed by linking together segments in the 2D radar elevation scans. The

height_of_the_BWER is the distance between the lowest candidate and the highest one, that is, the vertical length of the BWER. The best_height is the height of the highest candidate. The heights are computed using a 4/3 earth formula, assuming the standard atmospheric propagation, and are the height above the radar level. The VIL value at a location is the sum of all observed radar reflectivities (converted to liquid water content) in a vertical column above this location. The VIL at a pixel is, thus, an integration over all elevation scans. The low_VIL is the VIL computed from all elevation scans up to the elevation scan that contains the BWER.

The last three features are prev_conf, cov_vol, and sweep. The prev_conf is a measure of confidence of the closest candidate in the previous volume scan. It is computed by taking the current spatial location of the BWER, searching the list of candidates from the previous volume scan, and finding that closest to the BWER. If the distance of the closest match is within a few kilometers, then its confidence is used as prev_conf. The cov_vol and sweep features are the volume scan and sweep in which BWER was found.

These features were used by Lakshmanan and Witt (1996) and Lakshmanan (2000) also. To select a set of useful features, we ranked these features based on standard deviation. Features with very low standard deviations are discarded. In addition to this, we also used the online feature selection method of Pal and Chintalapudi (1997) to select/rank features. The top-ranked features suggested by this method were considered to be useful by experts also. We experimented with different sets of top-ranked features and found that the top seven features as ranked by the online feature selection method (Pal and Chintalapudi 1997) are adequate, and the use of more than seven top-ranked features does not improve the performance. Hence we settled for seven features.

The set of selected features are the minimum_rfl, average_rfl, maximum_bound, average_bound, average_cap, height_of_the_BWER, and low_VIL.

Each feature is then linearly normalized between [0, 1]. In other words, each input x is normalized by the formula $(x - x_{min})/(x_{max} - x_{min})$, where x_{max} and x_{min} are the maximum and minimum values of x.

b. Generation of fuzzy rule base

We use exploratory data analysis techniques to design the initial rule base (Delgado et al. 1997; Bezdek et al. 2005; Pal et al. 1997, 2002).

Let $\mathbf{X}^{Tr} = \mathbf{X}_1^{Tr} \cup \mathbf{X}_2^{Tr} \cup \mathbf{X}_3^{Tr}, \mathbf{X}_i^{Tr} \cap \mathbf{X}_j^{Tr} = \phi, i \neq j$, $i, j = 1, 2, 3$ be the training dataset, where $\mathbf{X}_i^{Tr}$ is the training dataset corresponding to Class$_i$. The three classes are

$$Class_i = \begin{cases} \text{Strong BWER,} & i = 1 \\ \text{Marginal BWER,} & i = 2 . \\ \text{No BWER,} & i = 3 \end{cases} \quad (1)$$

Now we apply the fuzzy c means (FCM) (Bezdek et al. 2005) clustering algorithm on each $\mathbf{X}_K^{Tr} \subset \mathbf{R}^p$, where K = 1, 2, 3, to cluster it into C_K clusters. In the present case the number of features $p = 7$.

In the FCM algorithm a cluster is represented by a centroid and the algorithm finds the centroids minimizing a sum of weighted squared distances of the data points from the centroids. The weights are related to memberships of data points to different clusters.

For each $\mathbf{X}_K^{Tr}$, the FCM algorithm produces C_K centroids, $\boldsymbol{\mu}_j^K, j = 1, 2, \ldots, C_K, \boldsymbol{\mu}_i^K \in \mathbf{R}^p, K = 1, 2, 3$.

The ith cluster of $\mathbf{X}_K^{Tr}$, can be represented by a fuzzy rule of the form R_j^K: if $\mathbf{x}$ is close to $\boldsymbol{\mu}_j^K$ then $\mathbf{x}$ belongs to class K. In R_j^K and $\boldsymbol{\mu}_j^K$, the subscript indicates the cluster number (rule number) and the superscript corresponds to the class label. To illustrate the rule structure, for notational clarity, for the time being we ignore the subscript and superscript. Suppose $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^T \in R^p$ is a cluster center for the strong BWER class. The associated rule is R: if $\mathbf{x}$ is close to $\boldsymbol{\mu}$ then the class is strong BWER. The fuzzy set "$\mathbf{x}$ is close to $\boldsymbol{\mu}$" can be represented by a membership function, where the membership to the set is inversely related to the distance of $\mathbf{x}$ from $\boldsymbol{\mu}$ (Bezdek et al. 2005).

Here $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)^T$ is a feature vector computed from a segment of a radar scan image that needs to be classified. The antecedent clause "$\mathbf{x}$ is close to $\boldsymbol{\mu}$" can be represented by a set of seven simpler atomic clauses and the rule can be rewritten as follows: if x_1 is close to μ_1 and x_2 is close to μ_2 and $\ldots$ and x_6 is closest to μ_6 and x_7 is close to μ_7, then the class is strong BWER. In this way we can generate the initial set of fuzzy rules for each cluster in $\mathbf{X}_K^{Tr}$ for $K = 1, 2, 3$.

Assume that there are $M = \Sigma_{K=1}^3 C_K$ rules in the rule base. So there will be M membership functions for each feature. For the jth rule, we model the fuzzy set "x_i is close to μ_{ji}" (linguistic value) by a Gaussian membership function as defined in Eq. (2),

$$\beta_{ji}(x_i) = e^{\frac{-(x_i - \mu_{ji})^2}{\sigma_{ji}^2}} \quad \text{where} \quad \sigma_{ji} > 0. \quad (2)$$

Note that other choices, such as a triangular membership function, can also be used. The initial spread of the Gaussian function (σ) is computed as the standard deviation of the ith feature of the data points in the associated cluster of the training data. Without loss of generality, let us denote the rules as $R_j, j = 1, 2, \ldots, M$. Also, let us denote the set of rules for the three classes strong BWER, marginal BWER, and no BWER by R_S,

R_M, and R_N, respectively. This is done just for notational clarity so that we can avoid use of another index for indicating the classes.

Given $\mathbf{x} \in R^p$, to decide its class label we compute the firing strength $\alpha_j(\mathbf{x})$ of the jth rule using some T norm (Klir and Yuan 1995). Here we use minimum as the T norm. Thus,

$$\alpha_j(\mathbf{x}) = \min_i \{\beta_{ji}(x_i)\}. \tag{3}$$

Let $\alpha_l = \max\{\alpha_j(\mathbf{x})\}$ over all j, then $\mathbf{x}$ is assigned to the class of rule R_l. In other words, if rule $R_l \in R_S$, $\mathbf{x}$ is assigned to the strong BWER class; similarly, if $R_l \in R_M$, $\mathbf{x}$ is assigned to the marginal BWER class, and if $R_l \in R_N$, $\mathbf{x}$ is assigned to the no BWER class; here $\{\}$ denotes a set.

We next provide a schematic description of the algorithm for generation of the initial rules: The input is $\mathbf{X} = \mathbf{X}_1^{\mathrm{Tr}} \cup \mathbf{X}_2^{\mathrm{Tr}} \cup \mathbf{X}_3^{\mathrm{Tr}}$. For each of $k = 1, 2$, and 3, five steps are performed: In step 1, cluster the data $\mathbf{X}_k^{\mathrm{Tr}}$ from class k to get C_k cluster centers $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_{C_k}$. In step 2, convert each cluster center $\boldsymbol{\mu}_i$ into a rule; if "$\mathbf{x}$ is close to $\boldsymbol{\mu}_i$" then the class is k. In step 3, rewrite the rule in step 2 as follows: If x_1 is close to μ_{i1} and $\ldots$ and x_7 is close to μ_{i7}, then the class is k. In step 4, model "close to μ_{ij}" by a Gaussian membership function; for example, x_1 is close to μ_{i1} by $\exp[-(x_1 - \mu_{i1})^2/\sigma_{i1}^2]$, where $\sigma_{i1} > 0$. In step 5, estimate the initial value of σ_{ij} by computing the standard deviation of the jth feature of the data points falling in the ith cluster of $\mathbf{X}_k^{\mathrm{Tr}}$. At the end of step 5, return to step 1 and repeat for the remaining k values.

c. Tuning of rule base

The tuning of the initial rule base is done by minimizing the training error using a gradient descent technique. In this regard we now derive the appropriate learning mechanism.

Let

$$\mathbf{x} \in \mathbf{X}^{\mathrm{Tr}}, \alpha_s = \max_{j, R_j \in R_S} \{\alpha_j\}, \alpha_m = \max_{j, R_j \in R_M} \{\alpha_j\} \quad \text{and}$$

$$\alpha_n = \max_{j, R_j \in R_N} \{\alpha_j\}.$$

In other words, α_s, α_m, and α_n are the maximum support provided by the rule base in favor of the three strong BWER, marginal BWER, and no BWER classes for $\mathbf{x}$. Let O_s, O_m, and O_n together define the class label

vector associated with $\mathbf{x}$. If the class assignments are crisp, then $O_s, O_m, O_n \in \{0, 1\}$; on the other hand, when the training data have fuzzy class labels, then $O_s, O_m, O_n \in [0, 1]$ where $[0, 1]$ indicates an interval from 0 to 1. If $O_s + O_m + O_n = 1$, then it represents a fuzzy label vector; otherwise, it is a possibilistic label vector (Bezdek et al. 2005). The error produced by the rule base on $\mathbf{x}$ is

$$E_x = (\alpha_s - O_s)^2 + (\alpha_m - O_m)^2 + (\alpha_n - O_n)^2. \tag{4}$$

If there are n data points in $\mathbf{X}^{\mathrm{Tr}}$, then the total error

$$E_{X^{Tr}} = \sum_{i=1}^{n} [(\alpha_s - O_s)^2 + (\alpha_m - O_m)^2 + (\alpha_n - O_n)^2]; \tag{5}$$

$E_{X^{Tr}}$ is a function of μ_{ji} and σ_{ji} of different rules that correspond to α_s, α_m, and α_n for different data points. To minimize $E_{X^{Tr}}$, we use gradient descent on the instantaneous error function E_x in Eq. (4). Thus, for every $\mathbf{x} \in \mathbf{X}^{\mathrm{Tr}}$ we compute the instantaneous error by the Eq. (4). Subsequently, we drop the subscript from E_x for the notational clarity.

We now update the membership parameters of the centroid (μ_j) and spread (σ_j) of the membership functions associated with the three rules as shown in Eqs. (6) and (7), respectively:

$$\mu_j(t + 1) = \mu_j(t) - \eta \frac{\partial E}{\partial \mu_j(t)} \quad \text{and} \tag{6}$$

$$\sigma_j(t + 1) = \sigma_j(t) - \eta \frac{\partial E}{\partial \sigma_j(t)}, \tag{7}$$

where $\eta > 0$ is the learning coefficient. The $\partial E/\partial \mu_j$ and $\partial E/\partial \sigma_j$ are derived in Eqs. (10) and (11). This may be viewed as refining the rules with respect to their contexts in the feature space.

Because $\min(x_1, x_2, \ldots, x_p)$ is not differentiable, we use a soft version of min in Eq. (8),

$$\mathrm{SoftMin}(x_1, x_2, \ldots, x_p) = \left(\frac{1}{p} \sum_{i=1}^{p} x_i^q\right)^{(1/q)}. \tag{8}$$

It is easy to see that SoftMin approaches min as $q \rightarrow -\infty$. Now, using Eqs. (2) and (8) in Eq. (5) we get Eq. (9),

$$E = \left\langle \left\{\frac{1}{p} \sum_{i=1}^{p} \left[e^{-\left(\frac{x_{is} - \mu_{is}}{\sigma_{is}}\right)^2}\right]^q\right\}^{(1/q)} - O_s \right\rangle^2 + \left\langle \left\{\frac{1}{p} \sum_{i=1}^{p} \left[e^{-\left(\frac{x_{im} - \mu_{im}}{\sigma_{im}}\right)^2}\right]^q\right\}^{(1/q)} - O_m \right\rangle^2$$

$$+ \left\langle \left\{\frac{1}{p} \sum_{i=1}^{p} \left[e^{-\left(\frac{x_{in} - \mu_{in}}{\sigma_{in}}\right)^2}\right]^q\right\}^{(1/q)} - O_n \right\rangle^2. \tag{9}$$

TABLE 1. Three typical rules before and after tuning.

Class	Tuning status	Rules
Strong BWER	Before	(0.77, 0.12), (0.78, 0.10), (0.58, 0.14), (0.66, 0.13), (0.49, 0.18), (0.40, 0.21), (0.65, 0.12)
	After	(0.70, 0.46), (0.78, 0.39), (0.27, 0.32), (0.48, 0.46), (0.12, 0.73), (0.36, 0.74), (0.64, 0.29)
Marginal BWER	Before	(0.62, 0.13), (0.64, 0.10), (0.63, 0.12), (0.60, 0.11), (0.20, 0.14), (0.50, 0.18), (0.46, 0.15)
	After	(0.24, 0.43), (0.33, 0.37), (0.76, 0.23), (0.63, 0.13), (0.29, 0.30), (0.55, 0.30), (0.38, 0.30)
No BWER	Before	(0.47, 0.12), (0.48, 0.11), (0.40, 0.14), (0.31, 0.11), (0.15, 0.16), (0.17, 0.13), (0.47, 0.12)
	After	(0.51, 0.32), (0.53, 0.37), (0.47, 0.54), (0.35, 0.36), (0.30, 0.63), (0.31, 0.03), (0.50, 0.37)

Differentiating Eq. (9) with respect to μ and σ, for the strong BWER case, we get

$$\frac{\partial E}{\partial \mu_{js}} = 2\left\langle \left\{\frac{1}{p}\sum_{i=1}^{p}\left[e^{-\left(\frac{x_{is}-\mu_{is}}{\sigma_{is}}\right)^2}\right]^q\right\}^{(1/q)} - O_s\right\rangle \times \left(\frac{1}{p}\right)^{(1/q)} \times 2\left\{\sum_{i=1}^{p}\left[e^{-\left(\frac{x_{is}-\mu_{is}}{\sigma_{is}}\right)^2}\right]^q\right\}^{(1/q)-1} \times e^{-q\left(\frac{x_{js}-\mu_{js}}{\sigma_{js}}\right)^2}$$
$$\times \frac{x_{js}-\mu_{js}}{\sigma_{js}^2} \tag{10}$$

and

$$\frac{\partial E}{\partial \sigma_{js}} = 2\left\langle \left\{\frac{1}{p}\sum_{i=1}^{p}\left[e^{-\left(\frac{x_{is}-\mu_{is}}{\sigma_{is}}\right)^2}\right]^q\right\}^{(1/q)} - O_s\right\rangle \times \left(\frac{1}{p}\right)^{(1/q)} \times 2\left\{\sum_{i=1}^{p}\left[e^{-\left(\frac{x_{is}-\mu_{is}}{\sigma_{is}}\right)^2}\right]^q\right\}^{(1/q)-1} \times e^{-q\left(\frac{x_{js}-\mu_{js}}{\sigma_{js}}\right)^2}$$
$$\times \frac{(x_{js}-\mu_{js})^2}{\sigma_{js}^3}. \tag{11}$$

Similarly, we can derive the learning rules for the marginal and no BWER cases. Note that for marginal and no BWER cases the subscript s will be replaced by m and n, respectively, in Eqs. (10) and (11).

The tuning process is continued till the change in error by an epoch becomes less than a predefined small threshold. In this way we get the final centroids (μ) and spreads (σ) of each fuzzy set associated with the rules. This rule set is now ready for classifying new data. The refined rule base is expected to result in a low error rate.

3. Results and discussions

In our investigation we have extracted four rules for each of the three classes. Out of the $4 \times 3 = 12$ rules, Table 1 shows only three typical rules—one for each class, before and after tuning. In Table 1, the "tuple" (μ, σ) indicates the center (μ) and spread (σ) of the Gaussian membership function.

A careful inspection of the rules before and after tuning in Table 1 reveals that tuning plays a significant role. For example, the rule corresponding to strong BWER (before tuning) is that if the minimum_rfl is close to 0.77 and average_rfl is close to 0.78 and . . . and low_VIL is close to 0.65, then the class is strong BWER.

The analysis of the rule base can also reveal very interesting characteristics about the underlying process. As an example, for the no BWER class the sixth feature, the height_of_the_BWER, is very important because the σ of the associated membership function is very small (0.03). Thus, the *specificity* of the fuzzy set, "the height_of_the_BWER is close to 0.31" is very high. This suggests that the vertical length (height of the suspected BWER) plays a very significant role in determining the no BWER case. *Such useful semantic information is clearly an advantage of a fuzzy rule–based system over other systems, such as neural networks.*

For this rule, the centers of Gaussian membership functions corresponding to only features 3, 4, and 5 changed significantly by the tuning; while the spread associated with each feature is changed significantly by the tuning to have a better coverage by the rule.

Now, we shall discuss the performance of our proposed model. The performance evaluation involves the use of a 2×2 contingency matrix **C**, defined in Eq. (12),

$$\mathbf{C} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{12}$$

In Eq. (12), d is the number of correctly detected (or forecasted) events or strong BWER (often referred to as *hits*), b is the number of false alarms (no BWER detected as strong BWER), c is the number of events not detected (*misses*), that is, strong BWER is detected

TABLE 2. Confusion matrix with training data (before tuning).

	Strong	Marginal	No
Strong	818	429	232
Marginal	151	460	124
No	1486	1755	1260

TABLE 3. Confusion matrix with test data (before tuning).

	Strong	Marginal	No
Strong	999	745	369
Marginal	161	350	131
No	1235	2159	1107

as no BWER, and a is the number of correctly classified nonevents; a is often difficult to estimate in the case of rare weather events like BWERs. The critical success index (CSI) is defined (Donaldson et al. 1975) using the contingency matrix as

$$CSI = \frac{d}{b + c + d}. \qquad (13)$$

We shall compute a, b, c, and d for two cases—(a) *before* and (b) *after* rule tuning.

a. Before rule tuning

Table 2 shows the confusion matrix for the training data before tuning the rules. The first row shows that out of total 1479 strong cases, 818 are classified as strong, 429 as marginal, and 232 as no BWER. We emphasize again, 1479 strong cases do not represent 1479 strong BWERs, but 1479 segments of a radar scan that are strongly suspected as part of some BWER. Similarly, the second row is for the marginal BWER and the third row for no BWER cases.

Using Table 2 we find $b = 1486$, $c = 232$, and $d = 818$.

Therefore, the contingency matrix (taking $a = 0$) corresponding to Table 2 is

$$\mathbf{C} = \begin{pmatrix} 0 & 1486 \\ 232 & 818 \end{pmatrix}, \qquad (14)$$

and CSI = (818/1486 + 232 + 818) = 0.32.

Similarly, Table 3 shows the confusion matrix with the test data before tuning the rules.

In this case, the contingency matrix is computed as

$$\mathbf{C} = \begin{pmatrix} 0 & 1235 \\ 369 & 999 \end{pmatrix} \quad and \quad CSI = 0.38.$$

b. After rule tuning

Table 4 shows the confusion matrix with the training data after tuning the rules. In comparing Table 4 with Table 2 we find a significant reduction in the number of mislabeled cases.

Table 4 results in the contingency matrix

$$\mathbf{C} = \begin{pmatrix} 0 & 764 \\ 174 & 972 \end{pmatrix} \quad and \quad CSI = 0.51.$$

Similarly, Table 5 displays the confusion matrix obtained on the test data after tuning the rules and the contingency matrix is

$$\mathbf{C} = \begin{pmatrix} 0 & 853 \\ 474 & 963 \end{pmatrix} \quad and \quad CSI = 0.42.$$

Lakshmanan (2000) used GAs to tune his bounded weak-echo regions detection algorithm. He reported results of five runs of the GA-based algorithm. In comparison with the average CSI of the five GA runs reported in Lakshmanan (2000), our results demonstrate a nearly 24% improvement in CSI on the training dataset and a 17% improvement in CSI on the test dataset.

4. Conclusions

The detection of a BWER signature using radar scans is often very difficult. The radar scans are first preprocessed to find the suspected regions. To generate the training data these suspected regions are manually labeled by experts. Each such region is represented by a set of features. Using exploratory data analysis an automatic scheme is developed for the extraction of fuzzy rules for classification of the suspected regions. It is a two-step process. In the first step we extracted an initial rule base using clustering. Then, in the second phase, we refined rules with respect to their context. In this regard, we derived the necessary update equations.

The proposed system is tested on a dataset not used in the training. Our system is found to produce an improved detection accuracy over the results reported in the literature on the same dataset. Our system extracts human-interpretable linguistic rules. We have demonstrated that such rules can reveal interesting information about the underlying process. This is a distinct advantage of the proposed system.

Our next step would be to integrate this information

TABLE 4. Confusion matrix with training data (after tuning).

	Strong	Marginal	No
Strong	972	333	174
Marginal	137	471	127
No	764	651	3086

TABLE 5. Confusion matrix with test data (after tuning).

	Strong	Marginal	No
Strong	963	676	474
Marginal	121	292	229
No	853	1497	2151

both spatially and temporally to find the location of each BWER. A genuine BWER should exhibit a cluster of suspected data points, and such clusters are expected to show up in successive radar scans. We plan to use the firing strength information to detect such clusters. We shall also investigate the use of cluster validity indexes to decide on the number of rules for each class. We expect that use of an appropriate feature set along with the right number of rules for each class will further improve the performance.

Acknowledgments. We sincerely acknowledge the editor and the anonymous reviewers for their valuable suggestions to improve the manuscript.

REFERENCES

Bezdek, J. C., J. Keller, R. Krishnapuram, and N. R. Pal, 2005: *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing.* Kluwer Academic, 776 pp.

Cotton, W. R., and R. A. Anthes, 1989: *Storm and Cloud Dynamics.* Academic Press, 883 pp.

Crum, T., and R. Alberty, 1993: The WSR-88D and the WSR-88D operational support facility. *Bull. Amer. Meteor. Soc.,* **74,** 1669–1687.

Delanoy, R. L., and S. W. Troxel, 1993: Machine intelligent gust front detection. *Lincoln Lab. J.,* **6,** 187–212.

Delgado, M., A. F. Gomez-Skorneta, and F. Martin, 1997: A fuzzy clustering-based rapid prototyping for fuzzy rule-based modelling. *IEEE Trans. Fuzzy Syst.,* **5,** 223–233.

Donaldson, R., R. Dyer, and M. Krauss, 1975: An objective evaluator of techniques for predicting severe weather events. Preprints, *Ninth Conf. on Severe Local Storms,* Norman, OK, Amer. Meteor. Soc., 321–326.

Dostalek, J. F., J. F. Weaver, and G. L. Phillips, 2004: Aspects of a tornadic left-moving thunderstorm of 25 May 1999. *Wea. Forecasting,* **19,** 614–626.

Howard, K., and J. Gourley, 1995: Evaluating the magnitude of uncertainties in WSR-88D measurements and measuring this impact on storm attribute trends—An initial study. National Severe Storms Laboratory Tech. Rep.

Klir, G. J., and B. Yuan, 1995: *Fuzzy Sets and Fuzzy Logic— Theory and Application.* Prentice Hall, 592 pp.

Krauss, T., and J. Marwitz, 1984: Precipitation processes within an Alberra supercell hailstorm. *J. Atmos. Sci.,* **41,** 1025–1034.

Lakshmanan, V., 2000: Using a genetic algorithm to tune a bounded weak echo region detection algorithm. *J. Appl. Meteor.,* **39,** 222–230.

——, and A. Witt, 1996: Detection of bounded weak echo regions in meteorological images. *Proc. IAPR'96,* Vienna, Austria, IAPR, 895–899.

Lemon, L. R., 1980: Severe storms radar identification techniques and warning criteria. National Severe Storms Forecast Center Tech. Rep. NWS NSSFC-3, 35 pp.

Markowski, P. M., 2002: Surface thermodynamic characteristics of hook echoes and rear-flank downdrafts. Part I: A review. *Mon. Wea. Rev.,* **130,** 852–876.

Pal, K., R. Mudi, and N. R. Pal, 2002: A new scheme for fuzzy rule based system identification and its application to self-tuning fuzzy controllers. *IEEE Trans. Syst. Man. Cybernetics B,* **32,** 470–482.

Pal, N. R., and K. K. Chintalapudi, 1997: A connectionist system for feature selection. *Neural Parallel Sci. Comput.,* **5,** 359–382.

——, K. Pal, J. C. Bezdek, and T. A. Runkler, 1997: Some issues in system identification using clustering. *Int. Joint Conf. Neural Networks: IJCNN 1997,* Piscataway, NJ, IEEE, 2524–2529.

Smalley, D., R. Donaldson, I. Harris, S. L. Tung, and P. Desrochers, 1995: Quantization of severe storm structures. *Proc. 27th Conf. on Radar Meteorology,* Vail, CO, Amer. Meteor. Soc., 617–619.

Smith, T., 1995: Visualization of WSR-88D data in 3-D using application visualization software. Preprints, *14th Conf. on Weather Analysis and Forecasting,* Dallas, TX, Amer. Meteor. Soc., 442–446.