1.5 Precipitation Classification Using Polarimetric Radar Data with A Support Vector Machine Method

Yadong Wang* and Lin Tang2,3

1Electrical and Computer Engineering Department, Southern Illinois University Edwardsville, Edwardsville, IL
2National Severe Storms Laboratory, Norman, OK
3Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK

ABSTRACT

A precipitation classification approach using support vector machine method is developed and tested on a C band polarimetric radar located in Taiwan (RCMK). Different from some existing classification methods that apply a whole volume radar data, the newly developed approach utilizes the unblocked polarimetric radar data from the lowest tilt to classify precipitation echoes into stratiform or convective types. In this approach, radar variables of reflectivity, differential reflectivity, the standard deviation of reflectivity, and the separation index calculated from the lowest tilt are utilized as the inputs, and the feature vector and weight vector in the support vector machine were optimized using well-classified training data. The proposed approach was tested with multiple precipitation events that include a mixture of widespread stratiform and convective, an isolated convective, and tropical convective precipitation. The results were compared with two existing approaches. The performance evaluation shows the proposed work can accurately identify the convective cells from stratiform precipitations with only the radar data from the lowest scanning tilt.

1. Introduction

Convective and stratiform precipitation systems exhibit significant differences in precipitation growth mechanisms and thermodynamic structures (e.g., Houghton 1968, Houze 1993, Houze 1997). Generally, convective precipitations are associated with strong and small areal vertical air motion (> 5 ms$^{-1}$), while stratiform types are associated with weak and mesoscale updrafts/downdrafts (< 3 ms$^{-1}$) (Penide et al. 2013). Moreover, a convective system generally consists of large and dense hydrometeors, which produces large high rainfall rate ($R$); stratiform precipitation on the other hand is associated with relative low $R$ (Anagnostou 2004). Accurately separating convective type precipitations from stratiform not only can promote the understanding of the cloud physics but also can enhance the performance of quantitative precipitation estimation (QPE) approaches. Numerous precipitation classification algorithms using ground in situ measurements or satellite were developed during the past four decades (e.g., Leary and Houze 1979; Tokay and Short 1996; Adler and Negri 1988; Hong et al. 1999).

Ground-based weather radar, such as Weather Surveillance Radar, 1988, Doppler (WSR-88D), can be used in severe weather detection, hydrometeor classification, QPE, and other applications. Precipitation classification methods were developed using single- or dual-polarization radars. For single-polarization radar, since only three moment variables in terms of reflectivity ($Z$), radial velocity ($v_r$), and spectrum width ($\sigma_v$) are available, the developed classification algorithms mainly rely on reflectivity and variables derived from it (e.g., Steiner et al. 1995, hereafter SHY95, Biggerstaff and Listemaa 2000; Anagnostou 2004, Yang et al. 2013, Powell et al. 2016). For example, based on the study from Steiner and Houze (1993), SHY95 proposed a widely used separation approach that utilizes the texture features derived from radar reflectivity field. Any grid point in $Z$ field with a value larger than 40 dBZ, or exceeds the average intensity taken over the surrounding background by specified thresholds is identified as the convective center. Those grid points surrounding the convective centers are classified as convective area, and far regions are classified as stratiform. During their experiments, Penide et al. (2013) found that SHY95 misclassified those isolated points embedded within stratiform precipitation and associated with low cloud-top height. Powell et al. (2016) modified the approach proposed by SHY95, and the new approach more effectively identifies shallow convection embedded within large stratiform regions and correctly identifies
isolated shallow and weak convolutions. A neural network convective-stratiform classification algorithm was developed by Anagnostou (2004), which utilizes six variables as inputs including storm height, reflectivity, vertical gradient of reflectivity, and other threes. Similar variables are also used in the fuzzy logic-based classification approach proposed by Yang et al. (2013). In these two approaches, a full volume of radar reflectivity field is needed to calculate variables such as the product of rain column top height reflectivity value at 2 km, and the vertically integrated liquid water content. However, according to U.S. radar operations centers (ROC), the WSR-88D radars are operated without updating a complete volume during each volume scan especially during precipitation events. In order to early capture the storm development for weather forecast and to obtain a more accurate precipitation estimation, new radar scanning schemes are designed to reorganize the updating order for a high frequency in low elevation heights and a less frequency for a higher height. These new schemes include, for example, the automated volume scan evaluation and termination (AVSET), supplemental adaptive intra-volume low-level scan (SAILS), the multiple elevation scan option for SAILS, and the mid-volume rescan of low-level elevations (MRLE). In the real-time operation, those classification results cannot be updated in a high frequency when data from a high tile is not available in time.

Polarimetric radars transmit and receive electromagnetic waves along the horizontal and vertical directions, and therefore can provide extra information about hydrometer’s size, shape, species, and orientation. A C-band polarimetric radar approach was developed by Bringi et al. 2009 (hereafter BAL), which classifies the precipitation into stratiform, convective and transition regions based on retrieved drop size distribution (DSD) characteristics. However, the performance of DSD-based approach depends on the environment regime (Thompson et al. 2015). Strong stratiform echoes might have similar DSDs to weak convective echoes, and lead to wrong classification results (Powell et al. 2016).

In this work, a novel precipitation classification algorithm was developed and tested on a C-band polarimetric radar located in Taiwan. This method classifies precipitations into stratiform and convective types using a support vector machine (SVM) method. Different from some existing classification approaches that utilize the whole volume of radar data, this new approach uses the unblocked data from the lowest scanning tilt. All the parameters used in the current approach are trained from typical convective and stratiform precipitation events. This paper is organized as follows: section 2 introduces the radar features and the radar data processing; section 3 proposes the classification using support vector machine; performance evaluation is shown in section 4, and a discussion and summary are given in section 5.

2. Method

2.1 Radars and Joss-Waldvogel Disdrometers

In the current work, the SVM precipitation classification approach was developed and validated on a C-band polarimetric radar (RCMK) located at Makung, Taiwan. The Weather Wing of the Chinese Air Force deployed this radar and made the data available to the Central Weather Bureau (CWB) of Taiwan since 2009. Together with other five single-polarization S-band WSR-88D (RCCG, RCKT, and RCHL) and one dual-polarization S-band radar (RCWF), these five radars provide real-time QPEs for CWB to support missions of flood monitoring and prediction, landslide forecasts and water resource management. Operating with a wavelength of 5.291 cm, RCMK performs volume scans of 10 tilts (0.5°, 1.4°, 2.4°, 3.4°, 4.3°, 6.0°, 9.9°, 14.6°, 19.5° and 25°) every 5 minutes with the range resolution of 500 m and angular sampling of 1 degree.

In order to understand drop size distribution (DSD) features from different types of precipitations, this work applies the DSD data of total 4306 minutes between 2013 to 2014 collected by four impact-type Joss-Waldvogel disdrometers (JWD). These four JWDs locate at the center of Taiwan as shown in Fig. 1. The measurement range and temporal resolution of these JWDs are 0.359 mm ~ 5.373 mm and 1 minute, respectively.

![Figure 1. The terrain of Taiwan, and the location of C-band polarimetric radar RCMK (marked with black square), JWDs (marked with black stars), and four S-band single polarization radar RCCG, RCKT, RCHL, and RCWF (marked with black circles).](image-url)

2.2 Input Polarimetric Radar Variables and Preprocess
The inputs of the SVM approach are four variables: reflectivity \( Z \), differential reflectivity fields \( Z_{DR} \), the standard deviation of \( Z \) \( \sigma_Z \), and separation index \( i \). In most of the classification approaches, \( Z \) is used as the input because the reflectivity field from convective systems generally shows higher values than from stratiform systems (e.g., Steiner et al. 1995, hereafter SHY95, Anagnostou 2004, Yang et al. 2013). In the approach developed by SHY95 and Powell et al. (2016), a radar echo with the reflectivity of 40 dBZ and above is automatically classified as convective type. The reflectivity value at 2-km elevation was used as one of the inputs in S-band classification (e.g., Anagnostou 2004, Yang et al. 2013). The second input viable is the differential reflectivity. Stratiform precipitations generally consist of condense of small to median raindrops, which generally produce a low value of \( Z_{DR} \). The convective precipitation, on the other hand, may produce large \( Z_{DR} \) because they consist of large and oblate raindrops.

For short wavelength radars such as those of C-band and X-band, the \( Z \) and \( Z_{DR} \) fields may be significantly attenuated when the radar beam goes through a heavy precipitation zone. Therefore, both \( Z \) and \( Z_{DR} \) fields need be corrected from attenuation before applied in the precipitation classification. Different attenuation correction methods were proposed using the differential phase \( \phi_{DP} \) measurement such as linearly \( \phi_{DP} \) approach, ZPHI method, and iterative ZPHI method (e.g., Jameson 1992, Carey et al. 2000, Testud et al. 2000, Park et al. 2005). Because of its simplicity and easy implementation in a real-time system, the linear \( \phi_{DP} \) method was applied in the current work.

\[
Z(r) = Z'(r) + \alpha (\phi_{DP}(r) - \phi_{DP}(0)) \quad (1.a)
\]
\[
Z_{DR}(r) = Z_{DR}'(r) + \beta (\phi_{DP}(r) - \phi_{DP}(0)) \quad (1.b)
\]

where \( Z'(r) (Z'_{DR}(r)) \) is the observed \( Z (Z_{DR}) \) at range \( r \), \( Z(r) (Z_{DR}(r)) \) is the corrected value; \( \phi_{DP}(0) \) is the system \( \phi_{DP} \) value; and \( \phi_{DP}(r) \) is the filtered (by FIR filter) differential phase. The attenuation correction coefficients \( \alpha \) and \( \beta \) depend on DSD, drop size shape relations (DSR), and temperature variations. The typical ratio of \( \alpha (\beta) \) is 0.06-0.08 (0.01-0.03) dB deg. \(^{-1}\) for C-band radars (e.g., Carey et al. 2000; Vulpiani et al. 2012). The \( Z \) and \( Z_{DR} \) fields are further smoothed with a 3 (azimuthal) by 3 (range) moving window function after corrected from attenuation.

Using the separation index \( i \) to separate convective precipitation from stratiform was originally proposed by BAL using a C-band polarimetric radar. According to BAL, \( i \) was proposed for a normalized gamma DSD:

\[
i = \log_{10}(N_w^{est}) - \log_{10}(N_w^{exp}) \quad (2)
\]
\[
\log_{10}(N_w^{exp}) = -1.6D_0 + 6.3 \quad (3)
\]

where \( N_w^{est} \) is the estimated \( N_w \) from observed \( Z_s \) and \( Z_{DR} \) as:

\[
N_w^{est} = Z/0.056D_0^{2.319} \quad (4)
\]
Support Vector Machines (SVM) Method

3.1 Introduction of SVM

Support vector machine (SVM) can be viewed as a kernel-based machine learning approach, which nonlinearly maps the data from input space to a high-dimension feature space followed by linearly mapping to the binary output space (e.g., Burges 1998). Given a set of training samples, the SVM constructs an optimal hyperplane, which maximizes the margin of separation between positive and negative examples (Haykin 2009). Specifically, given a set of training data \([X_i, y_i]_{i=1}^N\), the goal is to find the optimal weights vector \(W\) and a bias \(b\) such that:

\[
y_i(W^T X_i + b) \geq 1 \quad \text{for} \quad i = 1, 2, ..., N
\]

where \(x_i \in \mathbb{R}^m\) is the input vector, \(m\) is the input variable dimension (\(m = 4\) in this work), \(N\) is the number of training samples, and \(y_i\) is an output with the value of +1 or -1 that represents convective or stratiform, respectively. The particular data points \((X_i, y_i)\) for Eq. 6 is satisfied with the equality sign are called support vectors. The optimum weights vector \(W\) and a bias \(b\) can be obtained through solving the Lagrangian function with the minimum cost function (Haykin 2009). Since the SVM can be viewed as a kernel machine, finding the optimal weight vector and bias in Eq. 6 can be alternatively solved through the recursive least square estimations of:

\[
\sum_{i=1}^{N} \alpha_i y_i k(X, X_i) = 0
\]

With the solved \([\alpha_i]_{i=1}^N\), the SVM calculate the classification results with new input data \(Z \in \mathbb{R}^m\) as:

\[
f(Z) = \text{sign} \left( \sum_{i=1}^{N} \alpha_i y_i \phi(x_i)^T \phi(Z) \right)
\]

When \(f(Z) = 1\), the output is classified as convective, otherwise is classified as stratiform.

3.2 Training of the SVM

In SVM approach, the weight vector and bias in Eq. 6 need be optimized through a recursive least squares estimations using the training data set, and the training data play a critical role in the SVM approach. To ensure the training data are from right classified stratiform and convective types, the precipitation is first carefully examined following some general classification principles. For example, convective precipitation is associated with relative strong reflectivity, no obvious bight band signature, and high vertically integrated liquid (VIL). The precipitation type is then verified by ground observation, and then further confirmed by a multi-radar-multi-sensor (MRMS) precipitation classification approach implemented in Taiwan (Zhang et al. 2016). In this MRMS classification approach, the 3-dimensional radar reflectivity fields from 4 S-band single polarization radars (Fig. 1) are mosaiced. The composite reflectivity (CREF) together with other fields
such as temperature and moisture fields are used in the surface precipitation classification.

4. Performance Evaluation

4.1 Description of the experiments

The performance of the proposed approach was validated with three precipitation events from 2009 and 2011. Two experiments based on the BAL approach with different thresholds (i.e., BAL\textsuperscript{0} and BAL\textsuperscript{-0.5}) were also carried out in the current work. In these two experiments, the separation index $i$ from each pixel is first calculated using Eqs. 2~5 and a threshold of $T_0 = 0$ ($T_0 = -0.5$) is then used to separate convective type from stratiform types: the pixel is classified as convective when $i$ is larger than $T_0$; otherwise, it is stratiform. This work aims at a complementary method using separation index $i$ together with other variables to separate convective precipitation from stratiform type. Other classification approaches, introduced in section 1, use the data from multiple elevation angles, and their performance is not examined in this work.

The classification results from MRMS are used as the reference “ground truth” in the current work. Because the MRMS results are from the mosaic of four S-band single-polarization radars, the coverage and time stamp of the ground truth are off from the result of the single radar RCMK. The time difference between classification results from RCMK and MRMS could be as large as 5 minutes. Given the fact that the convective storm size, intensity, and cells locations could change during the 5-minute period, the pixel-to-pixel possibility of detection (POD) and false alarm rate (FAR) is not feasible to evaluate the performance. Therefore, we introduce a convective to a stratiform ratio ($R^{CS}$) to qualitatively evaluate the performance:

$$R^{CS} = \frac{N^{con}}{N^{str}}$$

(7)

Where $N^{con}$ and $N^{str}$ are the total pixel numbers of convective and stratiform types, respectively. More details about the evaluation results are shown in the following sections.

4.2 Experiment results

The performance of the proposed approach was first validated with two widespread mixture of stratiform and convective precipitation events from 30 August 2011 and 14 June 2012. For these two cases, 24-hour data (0000 UTC ~ 2400 UTC) was used in the evaluation. The results from BAL approach (BAL\textsuperscript{0} and BAL\textsuperscript{-0.5}) with different thresholds of 0 and -0.5 were also calculated. It should be noted that threshold of -0.5 is lower than the value suggested by BAL, and more pixels will be classified as convective by BAL\textsuperscript{-0.5}. The classification results from proposed SVM were calculated using the trained weight vector and biases. The convective ratio from MRMS, SVM, BAL\textsuperscript{0} and BAL\textsuperscript{-0.5} were calculated using Equation 10.

The time series plots of $R^{CS}$ are shown in Figure 3, where results from 30 August 2011 and 14 June 2012 are shown on A and B, and the $R^{CS}$ from MRMS, SVM, BAL\textsuperscript{0} and BAL\textsuperscript{-0.5} are presented by thick solid, thick dashed, thin solid and thin dashed lines, respectively. In general, BAL\textsuperscript{-0.5} classifies more pixels as convective than BAL\textsuperscript{0} as expected for both cases, and SVM shows the most similar results to MRMS comparing to BAL approaches. For the 30 August 2011 case (Figure 3A), if the MRMS result is viewed as the ground truth, BAL\textsuperscript{0} shows obvious under classification of convective type during this 24-hour period, but BAL\textsuperscript{-0.5} shows better performance. On the other hand, BAL\textsuperscript{-0.5} classifies more pixels as convective type than MRMS in the 14 June 2012 case (Figure 3B), but the results from BAL\textsuperscript{0} are more consistent with MRMS outputs. The overall $R^{CS}$ from MRMS, SVM, BAL\textsuperscript{0} and BAL\textsuperscript{-0.5} are shown in Table 1.

To better understand the performance of each approach, the classification results and radar variables ($Z$, $Z_{DR}$, and $i$) from two distinct moments were examined and shown in Figures 4~7. Classification results from 03:03 UTC 30 August 2011 were first shown in Figure 4, where MRMS, SVM, BAL\textsuperscript{0} and BAL\textsuperscript{-0.5} are shown in panel A, B, C, and D, respectively. The time stamp for MRMS result is 03:00 UTC, and the time difference from other three approaches is about 3 minutes. The three input variables of SVM at 03:03 UTC are shown in Figure 5, where $Z$, $Z_{DR}$, and $i$ are presented in panel A, B, and C. From Figures 3 and 4, it could be found that the $R^{CS}$ from MRMS, SVM, and BAL\textsuperscript{-0.5} show similar value, but $R^{CS}$ from BAL\textsuperscript{0} is obviously low. Within the black circle of Figure 5, the averages of $Z$ and $Z_{DR}$ both show relative large values ($Z > 36$ dBZ and $Z_{DR} > 0.75$ dB), this is a clear indication of convective type precipitation. Both SVM and BAL\textsuperscript{-0.5} classify most of the area within the black circle as convective, and this result is consistent with the MRMS result. Since the separation indexes within the black circle are below or slight higher than 0, most of the area is classified as stratiform type. For this moment, threshold -0.5 shows better performance than 0.
Figure 4. The classification results from BAL\(^0\) (A), BAL\(^{-0.5}\) (B), SVM (C) and MRMS (D). The time stamp for BAL\(^0\), BAL\(^{-0.5}\), and SVM is 03:03 UTC 30 August 2011, and time stamp for MRMS is 03:00 UTC 30 August 2011.

Figure 5. Radar variables of reflectivity (A), differential reflectivity (B), and separation index (C). The radar data was collected by RCMK at 03:03 UTC 30 August 2011.
5. Summary and Conclusion

A novel precipitation classification approach using support vector machine approach was developed and tested on a C band polarimetric radar located in Taiwan. Different from some existing classification algorithms that use whole volume scan data, the proposed approach only utilizes the data from the lowest unblocked tilt to separate precipitation into convective and stratiform types. A support vector machine method is used to integrate four inputs in terms of reflectivity, differential reflectivity, separation index, and standard deviation of reflectivity. The weighing vector and bias used in the support vector machine were trained with typical stratiform and convective precipitation events. Comparing to another polarimetric radar based approach developed by BAL, the proposed approach shows better results during three testing cases.

References


