Operational Utility of Dual-Polarization Variables in Lightning Initiation Forecasting

CRYSTAL J. WOODARD*, LAURENCE D. CAREY**, WALTER A. PETERSEN*, and WILLIAM P. ROEDER*

*UAHuntsville, Huntsville, Alabama
**NASA GSFC/WFF, Wallops Island, VA
^45th Weather Squadron, Patrick AFB, FL

(Manuscript received 16 January 2012; in final form 31 July 2012)

ABSTRACT

Radar-based operational forecasting algorithms for lightning initiation have relied heavily on the use of radar reflectivity. Surprisingly few, if any, algorithms rely on the use of variables specific to dual-polarization (DP) radar. A series of algorithms are synthesized based on past research and tested herein to determine if DP observations can improve lightning initiation forecasting. DP radar observations and a fuzzy-logic particle identification (PID) algorithm are used to identify the first instance of hydrometeors that are required for significant charging (e.g., graupel, hail, or their antecedent particles such as supercooled raindrops). Fifty ordinary convective cells are analyzed using DP observations. The potential for identifying lightning initiation is evaluated against a reflectivity benchmark. Although a wide range of reflectivity techniques are tested herein, the selected benchmark, 40 dBZ at –10°C, is well supported by many studies as a common, high-performing algorithm and is similar to operational algorithms currently in use. The benchmark results in 10.5 minutes of lead time with a perfect POD (1.0), 0.2 FAR, and 0.79 CSI. In comparison, the best first flash forecasting algorithm overall is another reflectivity based algorithm tested in this study, 40 dBZ at –15°C. The best algorithm produces perfect scores for POD, FAR and CSI with a reduced lead time of 8 minutes. Compared to the benchmark, DP variables result in decreasing FAR, as seen in the leading PID algorithm, detection of graupel at –15°C. The graupel algorithm results in a 0.93 POD, reduced FAR of 0.03, CSI of 0.91, and lead time of 9.5 minutes. The best differential reflectivity (ZDR) based algorithm produces a 0.93 POD, 0.09 FAR, 0.85 CSI, and a lead time of 11 minutes. Although the use of DP variables does not significantly improve skill (CSI) or lead time relative to the benchmark, it does reduce FAR by as much as 50%. More research is required to evaluate these small benefits relative to the increased cost in complexity and to verify the generality of the results.
1. Introduction

Lightning is a significant source of weather-related casualties and is currently the third leading cause of storm deaths, resulting in approximately 55 fatalities per year on average (Roeder 2012). Due to the factor of surprise, the first cloud-to-ground (CG) flash may be the most important to forecast, as it is likely the cause of many lightning related casualties. The onset of intra-cloud (IC) lightning aloft, which typically precedes the first ground flash (Williams et al. 1989), may be equally as important to forecast, providing valuable extra minutes of lead time for people to seek safety. Most lightning-related casualties are incurred during leisure and industry outdoor activities. Lightning nowcasting with longer lead times could provide officials at large outdoor venues with more time to respond to possible threatening weather events.

The ability to accurately forecast total (IC+CG) lightning, not just CG lightning, is also critical to aviation and space vehicle launch operations. For example, the 45th Weather Squadron (45 WS) issues lightning watches and warnings for Cape Canaveral Air Force Station, NASA Kennedy Space Center, Patrick Air Force Base, and other locations (McNamara et al. 2010). Forecasting natural lightning is also part of forecasting Lightning Launch Commit Criteria to protect in-flight space launch vehicles from natural and rocket-triggered lightning that could cause billions of dollars of damage and possible loss of life (McNamara et al. 2010).

To curb the loss of life and property damage, lightning initiation forecasting has been a focus of radar studies since the early years of radar operations and research. Many studies have examined radar reflectivity as a potential means for predicting the occurrence of lightning (Reynolds and Brook 1956; Larson and Stansbury 1974; Dye et al. 1986, 1989; Buechler and Goodman 1990; Michimoto 1991; Petersen et al. 1996; Roeder and Pinder 1998; Gremillion and Orville 1999; Vincent et al. 2003; Wolf 2006; Yang and King 2010; Mosier et al. 2011; Roeder and McNamara 2011). As a result, these reflectivity-based lightning forecasting techniques have been available to operational users for some time. Dual-polarization (DP) radar observations provide a unique ability to improve thunderstorm remote sensing via the identification of hydrometeor type, phase, shape, and concentration. Importantly, DP radar can assist in the determination of some of the precursor microphysical conditions relevant to lightning production and thus first flash forecasting. As part of a national plan, the NEXRAD (Next-Generation Radar) network of Weather Surveillance Radar-1988 Doppler (WSR-88D) radars, as well as many media-outlet owned radars, are being updated and converted to DP technology. Similarly, the 45WS has recently installed a C-band DP radar near Cape Canaveral Air Force Station (CCAFS) and KSC (Roeder et al. 2009). To our knowledge, no study has addressed, thus far, the use of DP radar variables on an operational basis for first flash forecasting.

As such, this study develops and tests operational lightning initiation nowcasting algorithms that use the full suite of DP radar variables and that are based on established relationships in the peer-reviewed literature. Section 2 summarizes key storm charging mechanisms related to lighting production, defines and describes DP radar variables, explains how DP radar can be used to identify these physical precursor conditions for significant charging and lighting production, and synthesizes this knowledge into some general classes of DP radar algorithms for first flash forecast testing. Section 3 describes the radar and lightning data used for this study and explains how the data are processed and the algorithms are tested on 50 convective cells over northern Alabama. Study results and associated discussion are presented in Section 4 while conclusions and suggestions for future work can be found in Section 5.
2. Dual-polarization radar-based lightning forecasting

Lightning initiation results from the charging mechanisms and microphysical processes that occur in a dynamic thunderstorm environment. A widely accepted method of thunderstorm charge accumulation, leading to lightning discharge, is the non-inductive charging (NIC) mechanism (Reynolds et al. 1957; Takahashi 1978; Dye et al. 1986, 1988; Saunders et al. 1991). At the microphysical level of particle-to-particle NIC, two ice particles of different sizes, such as graupel and an ice crystal, collide and separate net charge. The sign and magnitude of the charge separated by each particle are dependent upon environmental factors such as temperature and liquid water content (LWC; Takahashi 1978; Jayaratne et al. 1983; Saunders et al. 1991; Brooks and Saunders 1995). On the storm scale, dynamical processes (e.g., gravity and the updraft) effectively separate the charged hydrometeors into layered regions. In a general sense, charge separation develops a tripole structure in the thunderstorm (MacGorman and Rust 1998). The tripole structure consists of a main upper positive charge center over a main negative charge center that exists over a smaller lower positive charge center (Marshall and Rust 1991; Ziegler and MacGorman 1994; MacGorman and Rust 1998). The lower positive charge exists approximately below the –10°C height, and the main negative charge layer extends from –10°C to –20°C. Finally, the upper main positive charge approximately extends from the height of –20°C and above. These opposing charge layers are a key factor in the initiation of lightning through the development of a strong electric field, leading to atmospheric breakdown (MacGorman and Rust 1998). The region between –10°C and –20°C is the approximate location of the key hydrometeors and processes involved in particle- and storm-scale NIC and is therefore a focus of this research. Thus, to predict the possibility of lightning, radar is used to detect the presence of hydrometeors relevant to NIC, such as graupel, hail, or their antecedent particles (e.g., supercooled raindrops) within this important temperature zone of developing convection.

The DP radar provides estimates of the horizontal reflectivity ($Z_H$), differential reflectivity ($Z_{DR}$), the correlation coefficient ($\rho_{HV}$), and the specific differential phase ($K_{DP}$). From these observations, one can infer limited information about hydrometeor type, shape, size, concentration, and phase in a radar volume. One of the key variables of DP radar-based techniques is $Z_H$, which is often called simply “radar reflectivity” and is also measured by single-polarization radars. Radar reflectivity, which can be approximated by the 6th moment of the particle size distribution for Rayleigh scattering conditions (Bringi and Chandrasekar 2001; Rinehart 2010), provides a gross estimate of the concentration and especially the size of hydrometeors present in the range resolution volume. For reflectivity-only based techniques, temperature is used to constrain inferences regarding hydrometeor type or phase. Previous studies have been widely conducted using $Z_H$ to roughly identify the location of millimeter-sized hydrometeors, such as graupel, at temperatures below freezing, to estimate the location of large ice hydrometeors associated with NIC. Research found a relation between the first flash of lightning and the occurrence of $Z_H > 35–45$ dBZ at temperatures of –10°C or colder (Dye et al. 1986, 1989; Buechler and Goodman 1990; Zipser and Lutz 1994; Petersen et al. 1996; Roeder and Pinder 1998; Gremillon and Orville 1999; Vincent et al. 2003; Wolf 2006; Yang and King 2010; Mosier et al. 2011; Roeder and McNamara 2011). More specifically, of all the research conducted, 40 dBZ at –10°C is the most commonly supported leading reflectivity forecasting algorithm (Dye et al. 1989; Buechler and Goodman 1990; Gremillon and Orville 1999; Vincent et al. 2003; Wolf 2006; Yang and King 2010). For comparison with past radar-based lightning...
forecasting studies, we varied the $Z_H$ thresholds (35, 40, 45 dBZ) applied at specific temperatures ($-10^\circ\text{C}$, $-15^\circ\text{C}$, $-20^\circ\text{C}$), as summarized in Table 1.

**Table 1.** Table of 27 lightning initiation algorithms tested on radar data within the Larsen area, or echo area of $Z_H > 30$ dBZ at $-10^\circ\text{C}$, of the convective cells of interest. The test column indicates the variables used for the test. The middle column contains the various parameters used to compose the algorithms. Each parameter is then paired with each threshold level found in the third column. A full list of National Center for Atmospheric Research (NCAR) particle identification (PID) types used in this study can be found in Table 2.

<table>
<thead>
<tr>
<th>Test</th>
<th>Threshold Parameters</th>
<th>Threshold Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z_H$ and temperature</td>
<td>35, 40, 45 dBZ</td>
<td>$-10^\circ\text{C}$, $-15^\circ\text{C}$, $-20^\circ\text{C}$</td>
</tr>
<tr>
<td>Minimum echo top height requirement</td>
<td>15, 20, 25 dBZ</td>
<td>8 km</td>
</tr>
<tr>
<td></td>
<td>(combined with Larsen area</td>
<td></td>
</tr>
<tr>
<td></td>
<td>requirement at $-10^\circ\text{C}$</td>
<td></td>
</tr>
<tr>
<td>$Z_{DR}$ and $Z_H$ with temperature</td>
<td>$&gt; 0.5$ dB, $&gt; 1$ dB</td>
<td>$-10^\circ\text{C}$, $-15^\circ\text{C}$, $-20^\circ\text{C}$</td>
</tr>
<tr>
<td>($Z_H &gt; 40$ dBZ) ($Z_{DR}$ column)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PID and temperature</td>
<td>PID = 8,9 (graupel, small hail)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 &lt; PID &gt; 13 (small ice)</td>
<td>$-10^\circ\text{C}$, $-15^\circ\text{C}$, $-20^\circ\text{C}$</td>
</tr>
<tr>
<td></td>
<td>PID = 7, 9, 14 (supercooled drops)</td>
<td></td>
</tr>
</tbody>
</table>

One of the foundational studies to associate enhanced reflectivity areas at specific levels with lightning activity is Larsen and Stansbury (1974). Larsen and Stansbury found the occurrence of lightning correlates well with radar indicated precipitation areas identified by regions of radar reflectivity thresholds ($\geq 23, 43, 63$ dBZ) at a specific level (7 km). Following Larsen and Stansbury, Marshall and Radhakant (1978) demonstrated that lightning rates were functionally related to the radar echo area of $Z_H > 30$ dBZ at 6 km, which they called the “Larsen area”. A Larsen area is identified by the presence of a radar reflectivity threshold at a defined temperature level in the atmosphere. For the purpose of this study, the Larsen area is defined by the echo area of $Z_H > 30$ dBZ at $-10^\circ\text{C}$ (e.g., Fig. 1), which is roughly equivalent to Marshall and Radhakant’s definition over northern Alabama during the warm season. As defined, the Larsen area is associated with significant precipitation at the lower boundary of the main negative charge region of a typical tripole thunderstorm (MacGorman and Rust 1998) and is therefore closely associated with lighting potential (Larsen and Stansbury 1974; Stansbury and Marshall
1978; Marshall and Radhakant 1978). These early studies were forerunners to lightning initiation research, and like many initial studies focused on a few select cases. As discussed below and in the next section, we use the Larsen area as the minimum requirement for significant electrification and then investigate specific radar characteristics within these regions before assessing lightning potential.

**Figure 1.** Plan Position Indicator (PPI) imagery of $Z_H$ (dBZ, left) and $Z_{DR}$ (dB, right) at 5.6° elevation angle (or about $-10^\circ$C temperature level) on 4 June 2010 at 1824 UTC. Color shaded as shown. The area indicated with a dashed line is the Larsen area. The Larsen area is defined as the reflectivity threshold ($\geq 30$ dBZ) at a thermal threshold ($-10^\circ$C). The Larsen area is determined by the estimated center of the reflectivity echo and radius required to encompass the majority of the threshold reflectivity gates. The area indicated with the solid line is an example of a $Z_{DR}$ column. The enhanced values of $Z_{DR}$, $\geq 1$ dB, associated with high values of $Z_H$, $\geq 40$ dBZ, at approximately $-10^\circ$C indicate oblate raindrops at temperatures below freezing.

A more recent method of lightning initiation forecasting using DP radar variables is the identification of hydrometeors (e.g., graupel/small hail vs. supercooled rain drops) using significant values of $Z_H$ in combination with differential reflectivity ($Z_{DR}$) relative to temperature. Differential reflectivity is the decibel difference between horizontal reflectivity ($Z_H$, dBZ) and vertical reflectivity ($Z_V$, dBZ), or $Z_{DR} = Z_H - Z_V$ (dB). Physically, $Z_{DR}$ is the reflectivity weighted measure of hydrometeor shape or oblateness (Bringi and Chandrasekar 2001; Rinehart 2010). A small cloud or drizzle droplet is relatively spherical and would hence possess a $Z_{DR}$ value of $\approx 0$ dB. The larger a drop grows the more oblate it will become. A 3-mm rain drop would have a $Z_{DR}$ value near 1.3 dB. A tumbling hail stone has a $Z_{DR}$ value near zero due to the apparent spherical nature. The identification of supercooled raindrops in a column above the freezing level using significant $Z_H$ ($\geq 40$ dBZ) and elevated $Z_{DR}$ ($\geq 0.5$–$1.0$ dB) (Table 1), or so called $Z_{DR}$ column (e.g., Figs. 1, 2a-b, 3a-b), has been used to anticipate first lightning in past studies (Goodman et al. 1988; Jameson et al. 1996; Bringi et al. 1997; Carey and Rutledge 2000). In warm-based clouds, the lofted supercooled rain drops freeze at temperatures (T) colder than $-10^\circ$C and quickly rime, becoming nearly instantaneous sources of small hail or graupel. If supported aloft in a mature updraft at $T < -10^\circ$C, the small hail actively charge via
rebounding collisions with ice crystals. Once started, this NIC process can produce significant electrification and lightning within minutes. As such, the $Z_{DR}$ column is considered a leading indicator of lightning potential in warm-based clouds, such as over northern Alabama (Goodman et al. 1988), and is therefore a distinct algorithm tested in this study (Table 1).

The term “$Z_{DR}$ column” originates from the most common research approach for studying this process using vertical cross-sections of a storm’s core (Fig. 2a-b). As we use an operational approach to lightning forecasting herein, the $Z_{DR}$ columns in this study were identified from a quasi-horizontal cross-section produced by the plan position indicator (PPI) radar scans discussed in the next section (e.g., Figs. 1, Figs. 3a-b). As the supercooled rain drops freeze, they become a key ingredient to the formation of graupel and small hail. The rapid reduction of elevated $Z_{DR}$ values to near zero, (e.g. Figs. 2b-d, Figs. 3b-d), is an indication of the freezing of the $Z_{DR}$ column associated with the production of large ice hydrometeors, such as graupel and hail. Since large ice particles are essential for NIC, the freezing signature in $Z_{DR}$ suggests that lightning production could be imminent. These signatures make $Z_{DR}$ a key DP variable in hydrometeor identification algorithms for lightning initiation forecasting (e.g., graupel, small hail or their antecedent hydrometeors of supercooled rain in warm-based clouds).

Two other DP variables, the correlation coefficient ($\rho_{HV}$) and the specific differential phase ($K_{DP}$), can assist in hydrometeor identification. $\rho_{HV}$ is the correlation, or similarity, between the radar signal returned in the horizontal and vertical channels (Bringi and Chandrasekar 2001; Rinehart 2010). $\rho_{HV}$ is dependent on the scattering properties of the sensed medium and is useful for distinguishing between a mixed-phased volume, consisting of hail and rain, and a volume consisting largely of rain drops (or hail) alone. A mixed-phased volume would produce a less correlated signal due to the differences in hydrometeor shape, size, and phase in the volume thus producing a lower $\rho_{HV}$ value. $K_{DP}$ is the range derivative, at a specific range, of the differential phase ($\phi_{DP} = \phi_H - \phi_V$) between the horizontal ($\phi_H$) and vertical ($\phi_V$) propagation phases. $\phi_{DP}$ is caused by the modification of the polarized radar signal’s speed of propagation through mediums of different precipitation water content and thermodynamic phases (Bringi and Chandrasekar 2001; Rinehart 2010). $K_{DP}$ is dependent on orientation, shape, size, number concentration and phase of the hydrometeors. This variable assists in hydrometeor identification through inferences of the physical state (i.e., liquid vs. ice) of the hydrometeors present in a radar volume due to phase dependent propagation speeds (Bringi and Chandrasekar 2001; Rinehart 2010). For example, rain drops, which are oblate with larger dimension in the horizontal, slow a horizontally-polarized wave more than a vertically-polarized wave and therefore produce a positive $\phi_{DP}$ through the medium and positive $K_{DP}$ at a specific range. Conversely, small hail are associated with near zero values of $K_{DP}$ because of their near spherical shape, lower dielectric, and tumbling. Both $\rho_{HV}$ and $K_{DP}$ are used in conjunction with other variables for more accurate hydrometeor identification.
Figure 2. North (N) – South (S) vertical cross-sections of horizontal reflectivity ($Z_H$, dBZ, contoured every 5 dB starting at 20 dBZ) and differential reflectivity ($Z_{DR}$, dB, color shaded as shown) at a) 181837 UTC, b) 182537 UTC, c) 182944 UTC, and d) 183644 UTC during the development of the thunderstorm cell 1 of case 4 June 2010. All heights are above ground level (AGL). The N-S vertical cross-sections in a) – d) are taken through the cell of interest depicted in Fig. 3a-d, respectively, as it grows. The environmental 0°C temperature level is depicted with a solid line at 4.1 km. The −10°C temperature level is at an elevation of about 6.1 km (not shown). The 4-panel display demonstrates a typical evolution of warm-based thunderstorms from developing-to-mature stage. During the early stages (a-b), the $Z_{DR}$ column ($Z_{DR} > 1$ dB) associated with supercooled rain drops extended vertically well above the 0°C height level up to −10°C and colder. The vertical and horizontal extent of the $Z_{DR}$ column began to decrease in c) as the supercooled rain drops froze. Graupel and small hail (e.g., $Z_H \geq 35–40$ dBZ and $Z_{DR} \approx 0$ dB) were increasingly prevalent in c) and d) as the rain drops froze and likely grew via riming (see also Figs. 3c-d). Closely following production of graupel and small hail aloft, first lightning flash occurred at 1837 UTC in d).
Table 2. The modified National Center for Atmospheric Research (NCAR) PID algorithm contains 17 bulk hydrometeor categories which include the indication of insects, clutter and second trip signals. Included here are the categories significant to this research.

<table>
<thead>
<tr>
<th>Category</th>
<th>Color</th>
<th>NCAR PID</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>![Yellow]</td>
<td>Hail</td>
</tr>
<tr>
<td>7</td>
<td>![Green]</td>
<td>Rain and Hail</td>
</tr>
<tr>
<td>8</td>
<td>![Green]</td>
<td>Graupel and Small Hail</td>
</tr>
<tr>
<td>9</td>
<td>![Green]</td>
<td>Graupel and Rain</td>
</tr>
<tr>
<td>10</td>
<td>![Cyan]</td>
<td>Dry Snow</td>
</tr>
<tr>
<td>11</td>
<td>![Blue]</td>
<td>Wet Snow</td>
</tr>
<tr>
<td>12</td>
<td>![Purple]</td>
<td>Ice Crystals</td>
</tr>
<tr>
<td>13</td>
<td>![Pink]</td>
<td>Irregular Ice Crystals</td>
</tr>
<tr>
<td>14</td>
<td>![Pink]</td>
<td>Super-cooled liquid drops</td>
</tr>
</tbody>
</table>

The four previously discussed DP variables each assist in the identification of particle types. Particle identification is often difficult in an operational setting because the operator is forced to combine information from several DP parameters in a timely fashion in order to make a decision. Hence, robust automated methods for particle identification utilizing the entire collection of DP radar variables augmented by environmental information such as the temperature profile have been developed. One such application is the National Center for Atmospheric Research (NCAR) particle identification (PID) algorithm (Vivekanandan et al. 1999), which has been modified for C-band (Deierling et al. 2008). The PID uses a fuzzy-logic process incorporating the ensemble of DP variables to estimate bulk hydrometeor type (Vivekanandan et al. 1999). The NCAR PID utilizes the DP radar variables $Z_H$, $Z_{DR}$, $\rho_{HV}$, and $K_{DP}$ and temperature to classify the most probable hydrometeor type, or types in the case of a mixture, sensed by the radar (Table 2). On occasion, the probability of one PID classification (e.g., supercooled rain) may be only slightly greater than another (e.g., supercooled rain and graupel), leading to increased potential error associated with the PID. To reduce error in this study, the NCAR PID output categories (Table 2) have been grouped into significant hydrometeor categories related to NIC, such as supercooled drops, graupel/small hail and small ice (Table 1). Additional details on DP variables, their characteristics in various hydrometeor types and their use in hydrometeor identification algorithms can be found in Bringi and Chandrasekar (2001).
Another important feature to consider in radar thunderstorm forecasting beyond hydrometeor identification is echo top height. Previous studies examined radar echo top to determine the vertical development of thunderstorms associated with significant charging and first lightning (Lhermitte and Krehbiel 1979; Dye et al. 1989; Petersen et al. 1996). An elevated radar echo top height is indicative of a strong updraft that is necessary for charge separation. For example, Dye et al.’s study of warm season thunderstorms in the mountains of New Mexico (1989) demonstrated that significant charging typically does not occur until radar precipitation signatures reach at least 8 km. Small weak convective cells with limited vertical development and low radar echo tops (< 8 km) rarely produce lightning (Lhermitte and Krehbiel 1979; Dye et al. 1986, 1989; Williams 1989; Petersen et al. 1996; Gremillion and Orville 1999; Carey and Rutledge 2000; Vincent et al. 2003). In addition to the radar-inferred presence of graupel or small hail, a second set of constraints for first flash forecasting may therefore be required to insure that the updraft is sufficiently strong to keep the large ice hydrometeors lofted long enough for significant NIC and lightning. Similar to past studies, we assume that radar echo top extending to at least 8 km, or about −20°C in typical midlatitude warm-season thunderstorms, is indicative of a vigorous updraft capable of supporting large ice. As such, this study examines a radar echo top algorithm, which checks for the presence of significant reflectivity (Z_H > 15, 20, 25 dBZ) at an elevation of at least 8 km (Table 1). The echo top constraint is accompanied by the presence of a Larsen Area or area of Z_H ≥ 30 dBZ at −10°C, which is the minimum reflectivity consistent with the presence of large ice, such as graupel, in the charging zone.

Utilizing these tools and methods, various algorithms are developed for operational first flash forecasting using the DP radar. DP radar signatures are evaluated to detect the presence of hydrometeors significant to NIC (i.e., graupel and hail) and their antecedents (supercooled drops). The underlying hypothesis is that DP variables will provide significant improvements to first flash forecast skill and lead time compared to reflectivity alone because of DP radar’s ability to differentiate hydrometeor types necessary for significant NIC and lightning. The following section describes the process of data selection and the methodology conducted in this study.

3. Data and methodology

DP radar data were collected for an ensemble of 31 thunderstorms and 19 non-thunderstorm cells from eight case dates as listed in Table 3. As this operational application of DP radar observations to lightning forecasting is the first of its kind, we are focusing on the definition of the methodology and the quality of the analyses rather than the quantity of cases. As it was for reflectivity-based lightning forecasting techniques, operational radar studies at a preliminary stage typically follow earlier basic research observations (Larsen and Stansbury 1974, Dye et al. 1989) with carefully conducted manual case study analyses on relatively small samples (Gremillion and Orville 1999, Vincent et al. 2003) before applying the perfected techniques to larger samples (Wolf 2006, Mosier et al. 2011). For this DP study, 50 cells is considered the minimum sample size needed to conduct a relatively unbiased statistical analysis and is comparable to previous operational study sample sizes at the same preliminary stage (Gremillion and Orville 1999, Vincent et al. 2003).
Table 3. List of convective cell types for the eight case dates analyzed in this study. The first column indicates the date (year, month, day), the second column is the total number of convective cells selected from the case for examination, and the last column is the number of thunderstorm (T-Storm) cells.

<table>
<thead>
<tr>
<th>Case date</th>
<th>Total Cells</th>
<th>T-Storm cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>20100528</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>20100601</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>20100602</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>20100604</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20100729</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>20080411</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>20080625</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20080708</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

ARMOR, Advanced Radar for Meteorological and Operational Research, (Petersen et al. 2005, 2007) was used to collect the DP data for the 50 convective cells over the north Alabama region centered on Huntsville. The ARMOR DP radar operates using an optimized planned position indicator (PPI) scan mode for high temporal and spatial data resolution. High temporal and spatial resolution scanning of the mixed phased region is desired to capture the evolution of the microphysical processes and will affect sample quality and lead time results. To achieve quality data, the optimized temporal resolution of the sector volume scan (i.e., azimuthal extent of PPI’s < 360°) is 2 to 5 minutes. Furthermore, ARMOR is a C-band dual-polarization radar that is calibrated routinely, with data quality controlled and corrected for attenuation, differential attenuation, calibration and relative calibration errors (Zrnic and Ryzhkov 1996; Gorgucci et al. 1999; Carey et al. 2000; Bringi et al. 2001; Ryzhkov et al. 2005). After the data are quality controlled, cells of interest are visually located and assessed within a range of 15 to 100 km from the radar. A cell of interest is identified by adherence to a defined area of interest, or “Larsen area”. In other words, all thunderstorms meet the minimum requirements dictated by the Larsen area. As noted in Section 2, the Larsen area in this study is defined as 30 dBZ at −10°C. The storms evaluated are single-cell storms that are spatially distinct from any other nearby cells. These isolated storms are desirable in order to observe the development of precipitation relative to charging in a method similar to operational procedures while still ensuring the accuracy and quality of our assessments, such as lightning association.

Once a cell of interest is identified, the location of the first lightning flash (if any) is visually determined using data from the North Alabama Lightning Mapping Array (NA LMA),
which detects and locates IC and CG lightning (i.e., the total lightning flash count; Krehbiel et al. 2000; Koshak et al. 2004; Thomas et al. 2004; Goodman et al. 2005). Total lightning is significant to first flash forecasting because in general, IC lightning precedes CG lightning (Williams et al. 1989). In this study, the majority of CG lightning lags the occurrence of the first IC flash by an average of three minutes. However, it is possible for the first lightning in a thunderstorm to be a CG flash. For quality assurance, the first source of the flash is assessed by means of a VHF source-clustered flash algorithm (McCaul et al. 2005). Once the first flash is determined and the Larsen area defined, the ARMOR observations and processed PID data from the Larsen area are extracted to a file. The maximum data values from the Larsen area are evaluated with the test algorithms for each convective cell, including 31 observed thunderstorms and 19 observed non-thunderstorms. The list of 27 distinct test algorithms, which were outlined earlier in Section 2, is found in Table 1. The test algorithms evaluated in this study focus on the use of radar variables for the identification of hydrometeors, such as graupel, hail, and their antecedent, supercooled rain. The algorithms also incorporate significant temperature levels associated with the NIC mechanism of thunderstorm charging (e.g., the main charging levels of –10°C to –20°C). The algorithms include the use of radar reflectivity, as it suggests size and concentration of hydrometeors, reduced values of Z_{DR} implying drop freezing, radar reflectivity signatures of vertical development, and PID indications of significant hydrometeors. More details can be found in Section 2.

Lead time is determined based on the start time of the scan at the PPI tilt intersecting the level of interest during which the test algorithm thresholds are met and the first source of the first identified flash. This method reduces lead time errors to much less than a minute. To assess performance, the lead times and forecasting skill of each algorithm are compared to an established radar reflectivity benchmark, or 40 dBZ at –10°C, which is a common method in radar reflectivity-based lightning forecasting research and operations. The skill scores evaluated in this study are probability of detection (POD), false alarm ratio (FAR), critical skill index (CSI) and true skill statistic (TSS) (Jolliffe and Stephenson 2003).

Skill scores are a statistical analysis of forecast verification, or more plainly the ability of a forecast algorithm to correctly forecast or reject the occurrence of an event (Jolliffe and Stephenson 2003). These skill scores are derived from a 2×2 contingency table that evaluates DP radar-based forecasted events relative to actual observed occurrences. For each forecast algorithm, it is determined if the requirements of the algorithm are met for each convective cell. The evaluation of both thunderstorm and non-thunderstorm cells is required to complete a contingency table (e.g., hit – thunderstorm forecasted and occurred, miss – thunderstorm not forecasted and yet occurred, false alarm – thunderstorm forecasted and did not occur, and correct negative – thunderstorm not forecasted and did not occur) for all cells in the sample. A complete contingency table is required to accurately calculate POD, FAR, CSI and TSS (Jolliffe and Stephenson 2003).

4. Results and discussion

Lightning occurs during the mature phase of a thunderstorm after the development of significant ice precipitation in the mixed phase zone (–10°C to –40°C). The progression of a typical thunderstorm from this case study can be seen in a series of PPI images shown in Fig. 3 and the accompanied animation. These PPI figure panels display a cell, at approximately –10°C, from the first indication of a Larsen area leading up to the first lightning flash. The
accompanying vertical structures for Figs. 3a-d are shown in Figs. 2a-d, respectively. Initial development of the Larsen area is shown in Fig. 3a. At 1819 UTC, echo with $Z_H > 30$ dBZ aloft was associated with elevated $Z_{DR}$, suggesting supercooled raindrops at $T < 0^\circ$C (Figs. 2a, 3a). By 1826 UTC, elevated $Z_H$ and $Z_{DR}$ in the $Z_{DR}$ column (Figs. 2b, 3b) indicated large quantities of large supercooled rain drops, which was confirmed by the PID in Fig. 3b. By 1830 UTC, the freezing of the $Z_{DR}$ column had begun, as indicated by rapidly decreasing values of $Z_{DR}$ in Figs. 2c, 3c at $T < 0^\circ$C and especially $T < -10^\circ$C, where $Z_{DR}$ was already near zero and graupel was likely. The complete disappearance of the $Z_{DR}$ column (i.e., $Z_{DR} \approx 0$ dB) between 1830 UTC (Figs 2c, 2d) and 1837 UTC (Figs. 2d, 3d) supported the PID categorization of graupel and small hail at $T \leq -10^\circ$C during those times (Figs. 3c, 3d). Closely following the production of graupel at $T \leq -10^\circ$C, first flash occurred at 1837 UTC. The following sections will discuss the summary results for all 50 cells using each of the test algorithms conducted in this study by groups according to forecasting techniques (i.e., reflectivity based, $Z_{DR}$ column), and PID (Table 1).

a. Reflectivity based operational test algorithms

As discussed earlier in Sections 1 and 2, reflectivity based algorithms examining $Z_H$ thresholds associated with precipitation signatures at temperatures significant to the NIC method have been extensively studied. We find similar results in this study. The benchmark, 40 dBZ at $-10^\circ$C, results in an average lead time of 10.5 minutes with a perfect 1 in POD and moderate 0.2 FAR. These scores are associated with a CSI score of 0.79 and a TSS score of 0.58, which is only a moderate score. Decreasing the forecasting temperature threshold or increasing reflectivity requirements reduces FAR and effectively increases CSI scores (Fig. 4). The reduction of false alarm forecasts related to such reflectivity and temperature requirement changes also increases TSS scores. The reverse is true for lead time (Fig. 5) in that decreasing temperature or increasing reflectivity effectively decreases lead time. The decrease in lead time as a result of increased reflectivity values can be partly inferred from Fig. 3 and the associated animation. Figure 3a shows maximum values of 35 dBZ at 1818:37 UTC and less than two minutes later, as seen in the animation at 1820:04 UTC and again at 1825:37 UTC in Fig. 3b, values in excess of 40 dBZ are observed, effectively producing lower lead time for the latter threshold.

The trends in skill scores and lead times related to increasing reflectivity and decreasing temperature thresholds are expected results. At the $-20^\circ$C temperature threshold, FAR decreases significantly to values near zero with leading scores of CSI and TSS better than 0.90. However the lead times are significantly less with values as low as two and four minutes. More precisely, the lead times at $-20^\circ$C are nearly half of those associated with any other reflectivity based algorithm. While there is a desire for accurate skill, there is also a need for acceptable lead times. Based on the combined requirements for long lead time and high skill, the leading algorithm of basic reflectivity is 40 dBZ at $-15^\circ$C. This algorithm has perfect POD and FAR, resulting in leading CSI and TSS scores of 1 for the evaluated sample. However, it only produces a moderate lead time of 8 minutes, which is 2.5 minutes less than the benchmark.
Figure 3. $Z_H$ (dBZ), $Z_{DR}$ (dB) and PID evolution of perceptible growth in the thunderstorm cell 1 of case 6 June 2010 at approximately –10°C thermal level. UTC time of each scan: a) 181837, b) 182537, c) 182944, and d) 183644. Values are shaded as shown. In panel d the black star indicates the location of the first flash at 1837 UTC. See Table 2 for PID color key. From left to right, $Z_H$, $Z_{DR}$, and PID of cell 1 from initial detection of the Larsen area to first flash. Enhanced $Z_H$ values indicate perceptible sized particles key to the formation of the down draft. $Z_{DR}$ indicates the presence of oblate spheroids such as raindrops that contribute to precipitation growth. The PID confirms $Z_{DR}$ suggestion of supercooled water and graupel. Vertical cross-sections through cell in a-d) are shown in Figs. 2a-d, respectively. An animation of the time evolution of $Z_H$, $Z_{DR}$ and PID (from left to right) during the same time frame is shown here.
The presence of an 8 km or greater echo top based on various reflectivity thresholds at that height is tested as a forecasting tool to reduce FAR. Vertical development of a thunderstorm supports charge separation as related to NIC and the initiation of lightning. The echo height is not a standalone criterion as it must also accompany a base reflectivity threshold of \( \geq 30 \text{ dBZ} \) at \(-10^\circ\text{C}\), a minimum reflectivity signature suggesting the presence of large ice, such as graupel, in the charging zone. As seen in Fig. 4, all three 8-km reflectivity algorithms result in a 0.97 POD. Of the storms evaluated in this study, if the updraft is capable of lofting particles above 8 km with size and concentrations consistent with the presence of reflectivity equal to or exceeding 15 dBZ, it is also capable of lofting particles associated with 25 dBZ. Compared to the benchmark, the use of an 8-km minimum echo top height leads to a slight reduction of FAR resulting in values of 0.12 or less, thus leading to an increase in CSI skill by as much as 0.20 points. TSS skill associated with the use of an 8-km minimum echo top increases by as much as 0.28 in comparison to the benchmark. Increasing the reflectivity threshold used in defining minimum echo top at 8 km height reduces lead times. The best algorithm of this group is 30 dBZ at \(-10^\circ\text{C}\) and 15 dBZ above 8 km with the greatest lead time. The lead time of 10.5 minutes associated with this minimum echo top algorithm matches that of the benchmark. While the 8-km minimum echo top height algorithm does marginally reduce FAR, it also reduces POD. This leads to the conclusion that a single reflectivity threshold approach without consideration for echo top is still an effective algorithm. However, depending on the needs of the operational user, a minimum echo top height definition at 8 km might be found favorable.

**Figure 4.** Skill scores of the reflectivity threshold algorithms. Skill scores include probability of detection (POD), false alarm ratio (FAR), critical success index (CSI) and true skill statistic (TSS). The x-axis is the test algorithms, while the y-axis is the skill score value ranging from 0 to 1. A skill score for a particular algorithm with no corresponding color bar has a value of zero (0).
Figure 5. Lead times (minutes) of the reflectivity threshold algorithms. The average lead time is based on successfully forecasted events. The x-axis is the test algorithms, while the y-axis is the average lead time in minutes.

b. \(Z_{DR}\) column based operational test algorithms

A \(Z_{DR}\) column, as discussed previously (e.g., Figs 2a, 2b, 3a and 3b), is the location of supercooled drops lofted by and growing in the updraft identified by significant reflectivity values (e.g., \(Z_H > 40\) dBZ) and significant differential reflectivity values (e.g., \(Z_{DR} > 0.5\) dB). The use of \(Z_{DR}\) column algorithms at \(-10^\circ\)C, which is the same level as the benchmark, results in a very marginal half a minute increase in lead time (Fig. 6) and decrease in POD and FAR (Fig. 7). Considering \(Z_{DR}\)-based algorithms at \(-10^\circ\)C, the decrease of correctly forecasted events, which is seen in the reduced POD values of 0.97 and 0.93 in Fig. 7, is what leads to the slight increase in average lead times in Fig. 6. The sample size of this study is limited and the resulting difference of one to two cells incorrectly forecasted leads to the differences in the skill and lead time. The lead times decrease with a decrease in temperature, as expected and discussed previously. While the decrease in temperature also leads to reduced FAR, operational forecasters most concerned with accurate POD would likely not find the reduced FAR a favorable result at the expense of decreased POD. The difference in POD and FAR further leads to scores of 0.65 and 0.78 for TSS and 0.81 and 0.85 for CSI. Such values are only marginally better than the benchmark’s 0.58 TSS and 0.79 CSI scores. \(Z_{DR}\) column algorithms for temperature height levels greater than \(-10^\circ\)C decrease FAR at the expense of significantly reduced values of POD and lead time. The best algorithm in the \(Z_{DR}\) column based group is 40 dBZ or greater and \(Z_{DR}\) of 1 dB or greater at \(-10^\circ\)C. It is characterized by a low FAR of 0.09, a 0.93 POD and the leading TSS and CSI scores of the \(Z_{DR}\) column based group. While this algorithm reduces FAR by half, the lead time is only marginally increased at the expense of reducing POD with a marginal increase for CSI and TSS.
**Figure 6.** Lead times (minutes) of the $Z_{DR}$ column threshold algorithms. The average lead time is based on successfully forecasted events. The x-axis is the test algorithms, while the y-axis is the average lead time in minutes. The benchmark (40 dBZ at -10°C) lead time is shown to the left for comparison.

**Figure 7.** Skill scores of the $Z_{DR}$ column threshold algorithms. Skill scores include probability of detection (POD), false alarm ratio (FAR), critical success index (CSI) and true skill statistic (TSS). The x-axis is the test algorithms, while the y-axis is the skill score value ranging from 0 to 1. The benchmark results (40 dBZ at -10°C) are shown to the left for comparison. A skill score for a particular algorithm with no corresponding color bar has a value of zero (0).
c. **PID based operational test algorithms**

Automated algorithms can help the forecaster quickly and accurately identify hydrometeor types using the suite of dual-polarization variables (Section 2). Based on the requirements for NIC, algorithms for the identification of graupel, small ice particles and supercooled rain at significant levels are synthesized and tested. This method is hypothesized to provide the greatest increase in forecasting skill and lead time because it integrates all DP observables. The resulting scores of the PID algorithms can be seen in Fig. 8. Comparing the test algorithms at –10°C to the benchmark, the PID algorithms provide at least a minute increase in lead time (Fig. 9). A single minute increase in forecasting lead time relative to the benchmark is a marginal improvement, though it is the greatest increase of all the algorithms.

![Figure 8. Skill scores of the PID threshold algorithms. Skill scores include probability of detection (POD), false alarm ratio (FAR), critical success index (CSI) and true skill statistic (TSS). The x-axis is the test algorithms, while the y-axis is the skill score value ranging from 0 to 1. The PID categories correspond to the following ID numbers provided in Table 2: large ice – 8 and 9, small ice – 10 to 13, and supercooled [drops] – 7, 9, and 14. The benchmark results (40 dBZ at –10°C) are shown to the left for comparison. A skill score for a particular algorithm with no corresponding color bar has a value of zero (0).](image)

Additionally, FAR associated with the detection of PID graupel and supercooled drops decreases marginally at –10°C and more at increasing height levels to values as low as 0.03. Small ice PID algorithm performs very poorly in terms of total skill. Small ice results in perfect skill of POD with elevated FAR of 0.38, a value greater than even the benchmark. The increase in FAR is because many non-thunderstorm cells develop to a height of –10°C or greater with an updraft capable of developing and supporting small ice crystals. This algorithm could be found advantageous to some operations despite the increased FAR due to the increased lead times of 12.5 to 14.5 minutes, which are 2–4 minutes longer than the benchmark. Alternatively, we did not test a reflectivity threshold representative of small ice crystals; however previous studies
have examined similar low reflectivity values (Hondl and Eilts 1994). Hondl and Eilts found a high POD and FAR associated with lowered reflectivity threshold algorithms.

**Figure 9.** Lead time (minutes) results of the PID threshold algorithms. The average lead time is based on successfully forecasted events. The x-axis is the test algorithms, while the y-axis is the average lead time in minutes. The PID categories correspond to the following id numbers provided in Table 2: large ice – 8 and 9, small ice – 10 to 13, and supercooled [drops] – 7, 9, and 14. The benchmark (40 dBZ at −10°C) lead time is shown to the left for comparison.

The best algorithm of the PID category is graupel PID at −15°C based on a balance of skill and lead time. Graupel PID at −15°C results in a high POD of 0.93, a result of two inaccurately forecasted events. FAR is 0.03 which is much less than that of the benchmark’s 0.2 FAR. The reduced FAR leads to an increase in CSI, a result of 0.91, which is the leading scorer in this set. The TSS score for graupel PID at −15°C is 0.88, also the highest score in this set. The associated lead time is 9.5 minutes; which is a mere minute less than the benchmark.

5. Conclusions

The reflectivity based benchmark, 40 dBZ at −10°C, results in 10.5 minutes of lead time with a perfect POD (1.0), 0.2 FAR, and 0.79 CSI. Of all the tests conducted in this study, reflectivity based, ZDR column, and PID, the best result is from the reflectivity based algorithm, 40 dBZ at −15°C with perfect POD and FAR (0.0) and an 8-minute lead time. The best algorithm is determined based on skill scores, POD, FAR, CSI, and TSS, balanced with lead time. On the other hand, Air Force and other operational forecasters often have longer desired lead times. In consideration of this need, weighting lead time slightly more, the graupel PID at −15°C algorithm is a leading algorithm due to its longer lead time of 9.5 min, while still maintaining good performance with a high POD (0.93), reduced FAR of 0.03, 0.91 CSI, and 0.88 TSS. Furthermore the best ZDR column algorithm, which is 40 dBZ or greater and ZDR of 1 dB or
greater at \(-10^\circ\text{C}\), produces a 0.93 POD, 0.09 FAR, 0.85 CSI, and a lead time of 11 minutes. While DP variables generally decrease FAR, the increase in skill is only marginal as seen by CSI scores. Considering that the average lead time examined in this study derives from a limited sample size, a single minute or half a minute increase in lead time is a marginal increase. For this limited sample, DP variables are effective in reducing FAR but provide little advancement to lead time. While the reduction of FAR does result in a modest increase in general skill, such as CSI, it does not significantly improve upon current operational reflectivity based forecasting algorithms, as shown by comparison to the benchmark (e.g., CSI = 0.79, 0.85, and 0.91 for the benchmark of 40 dBZ at \(-10^\circ\text{C}\), the best $Z_{\text{DR}}$ column algorithm and graupel at \(-15^\circ\text{C}\), respectively).

The results of this study are based on a limited sample of ordinary single cell convection subjectively selected by the operator over northern Alabama. The expansion of the data set to include a larger sample size should reduce any possible bias error associated with a small sample set. A larger sample would assist in a more conclusive assessment of the potential advantages of DP variables. The data set herein was selected from a single geographical region. As seen in the results of reflectivity based forecasting algorithms, different geographical regions might produce different results. To mitigate this limitation, a follow-on study should be conducted in different regions to verify and extend the current study. In areas not covered by an LMA, the National Lightning Detection Network (NLDN) can be used to conduct a study examining CG lightning activity (Woodard 2011). An expanded sample should include different meteorological and convective situations (e.g., supercells, squall lines and mesoscale convective systems) for more robust results. This study examined ordinary convection so the lead times are reflective of short lived convective cells. Application of an automated algorithm to identify and track thunderstorm cells could provide a more unbiased result while enabling the analysis of a larger data set. Finally, a cost-benefit analysis of DP variables should be conducted through real-time application of DP algorithms by operational weather forecasters and users.

Acknowledgements. This work was funded at UAHuntsville by the NOAA/NWS Collaborative Science Technology and Applied Research (CSTAR) program (NA08NWS4680034) and the NASA Space Shuttle program and Terrestrial Environments Office at NASA MSFC under grants NNMO5AA22A and NNM11AA01A. We wish to acknowledge the assistance of Chris Schultz, Elise Schultz, Mariana Scott, Matthew Anderson and Patrick Gatlin during ARMOR radar operations and data analysis. We would like to acknowledge Dennis Buechler for his assistance with the NA LMA data. Finally, we thank Pete Wolf and an anonymous Reviewer for their thoughtful reviews and helpful suggestions.

REFERENCES


Reynolds, S. E., and M. Brook, 1956: Correlation of the initial electric field and the radar echo in thunderstorms. *J. Meteor.*, 13, 376–380.


