2.4

UTILIZING WDSS-II TO AUTOMATE DATASET PREPARATION FOR A STATISTICAL INVESTIGATION OF TOTAL LIGHTNING AND RADAR ECHOES WITHIN SEVERE AND NON-SEVERE STORMS

Scott D. Rudlosky * and Henry E. Fuelberg
Florida State University, Tallahassee, FL USA

1. Introduction

Cloud-to-ground (CG) lightning data from the National Lightning Detection Network (NLDN) have been available to National Weather Service (NWS) forecasters for over a decade. These data have been used for climatological studies (Fig. 1a) and for real time forecasting. A few NWS Weather Forecast Offices (WFOs) also receive total lightning data covering small regions within their county warning areas. These data come from local Lightning Mapping Arrays (LMAs) or Lightning Detection and Ranging (LDAR) networks (Fig. 1b-c). The upcoming GOES-R satellite will house a Geostationary Lightning Mapper (GLM) that will provide total lightning information to all NWS stations (Goodman et al. 2008).

Although CG and intracloud (IC) lightning are important in their own right, lightning also appears to be related to severe storms such as large hail producers, tornadoes, and damaging straight line winds. Figure 2 presents CG flash locations and preliminary severe storm reports for 5 May 2003 to illustrate the spatial collocation between CG lightning and severe storms. Total lightning data, used in conjunction with radar data, will provide researchers and forecasters with a better understanding of thunderstorm morphology and its relation to severe weather (Steiger et al. 2007). This suggests that lightning data can be used to assess the potential for severe weather events and to aid in warning decision support.

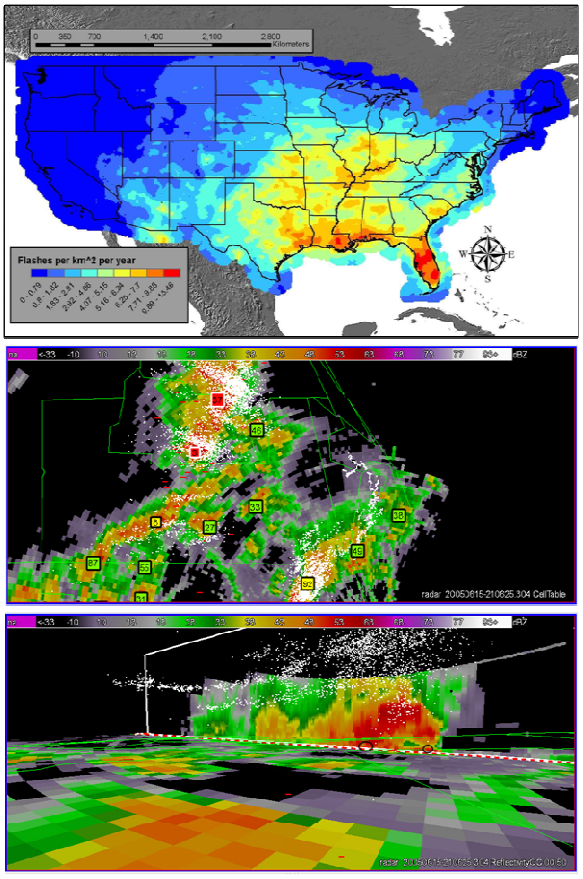

Fig. 1 a) Cloud-to-ground flash density from the NLDN computed on a 20 km grid between 2002-2006 b) Plan view of WSR-88D data, SCIT defined cells (boxes), and LDAR sources (white dots) for a 1 min period c) Cross-section during the same period as b) displaying the LDAR sources and the volume of reflectivity data.

Algorithms for ingesting and processing NLDN, LDAR/LMA, WSR-88D, and RUC-derived data are contained in the Warning Decision Support System

* Corresponding author address: Scott D. Rudlosky, Florida State University, Department of Meteorology, Tallahassee, FL 32306-4520 ; email: srudlosk@met.fsu.edu

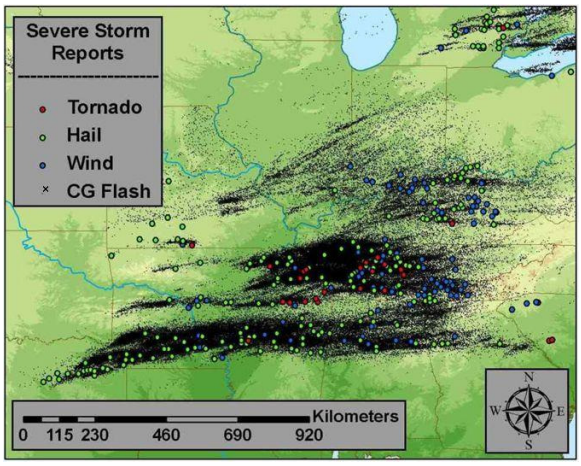

Fig. 2 NLDN cloud-to-ground flashes (black dots) and preliminary severe storm reports (colored circles) on 5 May 2003.

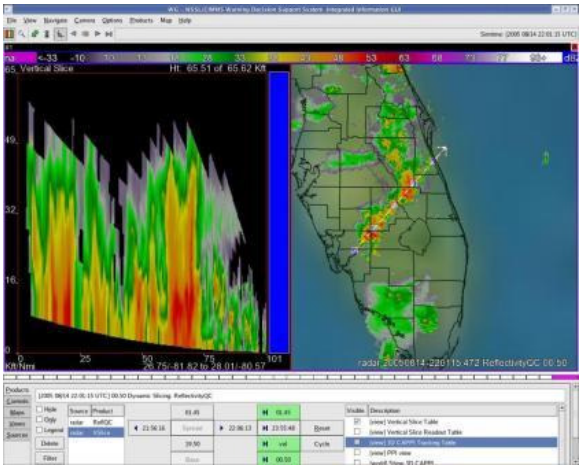

Fig. 3 Example of the graphical user interface (GUI) within the Warning Decision Support System – Integrated Information (WDSS-II) software. Both plan and cross-section views of quality controlled reflectivity data are shown.

– Integrated Information (WDSS-II, Lakshmanan et al. 2007) software.

These algorithms create spatial grids of meteorological fields, and can be modified to create higher-order, user-defined parameters. A major advantage of WDSS-II is that all data sources share a common spatial domain, allowing direct comparisons between them (Fig. 3). This paper describes our recent modifications to WDSS-II and illustrates the benefits of automating the development of the datasets needed to investigate lightning production within severe and non-severe storms.

Our overriding theme is to derive storm-scale relationships between lightning and radar data to ensure their best possible utilization in the NWS warning assessment process. Specifically, we seek to derive statistical relationships between radar-derived parameters and total lightning data in order to distinguish between severe and non-severe storms. These relationships then will be used to develop algorithms and guidelines as to whether a particular storm is likely to require a warning. Armed with these newly developed techniques, NWS offices will be able to better utilize current lightning data to assess severe storm potential and will be prepared for the deployment of the GLM.

2. Background

Numerous studies have investigated patterns of NLDN-derived CG lightning and their relation to the development of severe storms (Kane 1991, MacGorman and Burgess 1994, Knapp 1994, Biggar 2002, Carey et al. 2003). These studies focused on the CG flash rate, the location of CG flashes in both space and time, and the polarity of the CG flashes. For example, Biggar (2002) noted that a tornado touched down ~10 min after a significant polarity reversal from mainly positive CG (+CG) to mainly negative CG (-CG). His research suggested that CG data could improve tornado warning time by as much as 10 min for low-topped supercells that are distant from the radar. In a nationwide study, Carey et al. (2003) noted a positive trend between mean hail size up to 8 cm in diameter and the mean percentage of positive CG. Studies also have examined relations between the multiplicity and peak current of CG lightning and severe storms.

Total lightning signatures also have been related to severe storm development. Williams et al. (1999) found that sudden increases in total flash rate, denoted “lightning jumps”, preceded tornado occurrences in Florida by as much as 20 min. Lightning activity has been found to vary with updraft evolution, while updraft strength largely governs storm intensity (Gatlin 2007). Gatlin (2007) noted

that when lightning jumps alone are used as severe weather predictors, they produced a probability of detection of 0.985 and a false alarm rate of 0.446. The descending mesocyclone tornadoes in his study were preceded by lightning jumps by ~ 30 min, whereas the lead-time was 18 min for the non-descending type. Patterns in total lightning also have been related to severe hail and winds. Goodman et al. (2005) noted a strong increase in total lightning 9 min before damaging winds were observed at the surface during a severe summertime pulse thunderstorm.

Very few studies have compared the lightning characteristics of severe storms with those of non-severe storms. MacGorman and Morgenstern (1998) found differences between the percentages of +CG and –CG in twenty-five mesoscale convective systems compared to all thunderstorms in Oklahoma. Differences in peak currents and flash rates also were noted. Carey and Rutledge (2003) used NLDN data to investigate a region extending from the Kansas/Colorado border into Minnesota where positive anomalies in the percentages of +CG lightning and peak currents were observed. They found that the properties of lightning associated with severe storms during the warm season were different from those of non-severe storms. Montanya et al. (2007) noted that the average multiplicity of –CG flashes was 1.74 and 1.17 for severe and non-severe storms, respectively, while their average peak currents were -11.7 kA and -10.65 kA, respectively. Gatlin (2007) examined twenty storms over northern Alabama that produced severe storm reports. They then contrasted the total lightning characteristics of tornadic storms vs. those producing only large hail or damaging winds.

These types of studies should be expanded to include other regions of the country, and additional studies contrasting the lightning characteristics of severe vs. non-severe storms are needed for NWS forecasters to effectively use total lightning data in assessing severe storm potential.

Most previous research has defined storm cells using the Storm Cell Identification and Tracking

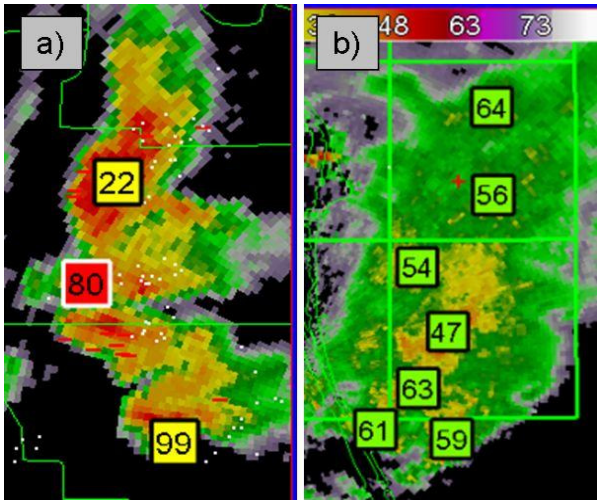

Fig. 4 Example of SCIT defined cells (numbered boxes) overlain on composite reflectivity. Also included are the LDAR flash initiation locations (white dots) and the cloud-to-ground flash locations.

(SCIT) algorithm (Fig. 4). Predefined radii linked the NLDN and LDAR/LMA data with individual storms. Although useful, SCIT tracking is intermittent and typically utilizes only a limited number of radar parameters. This has limited the scope of studies seeking to determine relationships between lightning, radar, and severe weather. However, the National Severe Storms Laboratory (NSSL) and others recently have developed new tools and techniques that facilitate these types of analyses. Newly developed software, enhanced data resources, and improved analysis techniques now allow detailed investigations of the relations between total lightning and radar within individual storms.

3. Our Methodology

The previously described relationships between total lightning, radar, and storm severity emphasize the need to quantitatively evaluate lightning data when accessing the potential for severe weather. Previously, the vast amounts of data and somewhat limited software capabilities necessitated a case study mode for storm analyses. However, seasonal and geographical variations strongly influence correlations between lightning and severe weather (Carey and Buffalo 2007). This suggests that robust statistics obtained from a large number of storms are

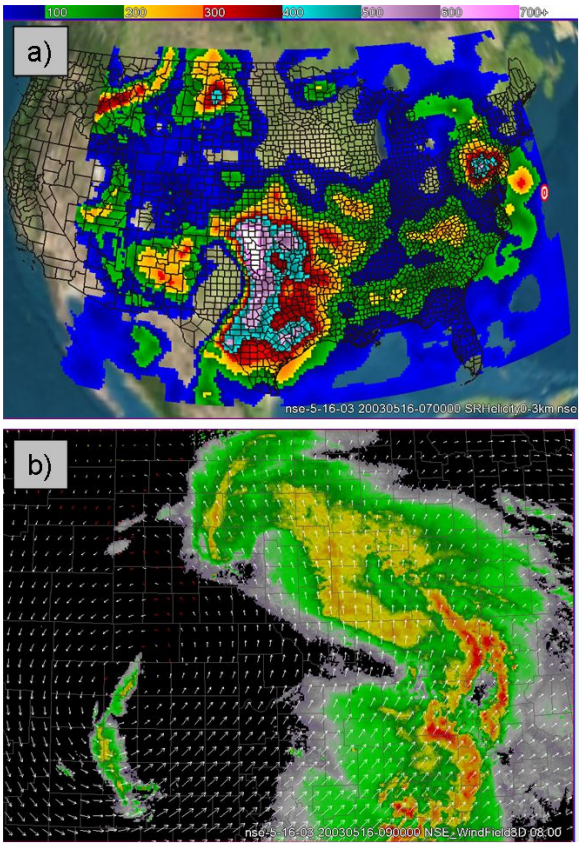

Fig. 5 a) RUC-derived storm relative helicity computed using the near-storm environment (NSE) algorithm within WDSS-II b) NSE winds and QC reflectivity at 8 km AGL.

needed to accurately quantify these relationships. Although developing such a dataset is not feasible using traditional methods, the WDSS-II software provides an ideal framework for developing lightning and radar databases and for making comparisons between them. This paper briefly describes WDSS-II and focuses on our algorithm modifications made in order to obtain a large database of storms from which to develop severe storm guidance.

WDSS-II allows users to simultaneously view and manipulate data from multiple sources. It contains a suite of algorithms to combine information about the Near-Storm Environment (NSE) from the Rapid Update Cycle (RUC) model with WSR-88D radar data to compute a

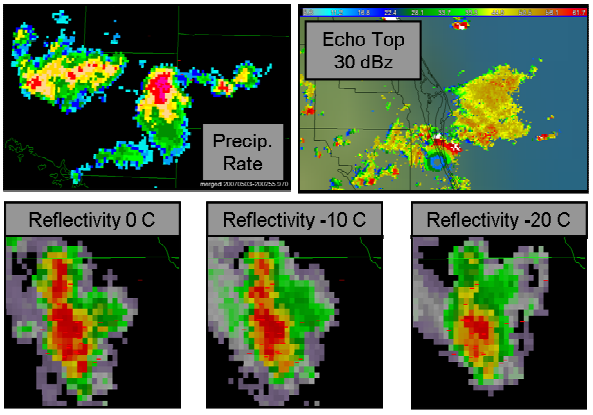

Fig. 6 Example fields resulting from the merger of the RUC near-storm environment (NSE) and WSR-88D parameters. These include the precipitation rate, the echo top at 30 dBZ, and the reflectivity at 0 ^{o}C, -10 ^{o}C, and -20 ^{o}C.

large number of desired parameters. For example, Fig. 5a displays RUC-derived storm relative helicity, while Fig. 5b illustrates our ability to overlay the RUC-derived parameters onto WSR-88D data. Specifically, Fig. 5b displays NSE winds overlain on reflectivity at 8 km AGL. The 3-D wind field is used to advect features and to compute higher order radar and RUC-derived parameters. These parameters range from basic reflectivity and radial velocity fields to more advanced parameters such as precipitation rates and rotation tracks. Figure 6 illustrates some example parameters resulting from the radar-RUC merger. These derived parameters then can be superimposed onto CG lightning information from the NLDN and/or total lightning data from LDAR/LMA networks to investigate lightning patterns within individual storms.

A major advantage to including lightning data is their temporal resolution. The improved resolution is used to enhance WSR-88D data to more accurately describe storm-scale processes. Various WDSS-II algorithms are modified and combined to create both basic and enhanced gridded fields of NLDN and LDAR/LMA parameters.

Characteristics of total, negative, and positive CG lightning are considered separately, and include the average polarity, multiplicity, and peak current within a specified grid cell (2×2 km). Currently, these

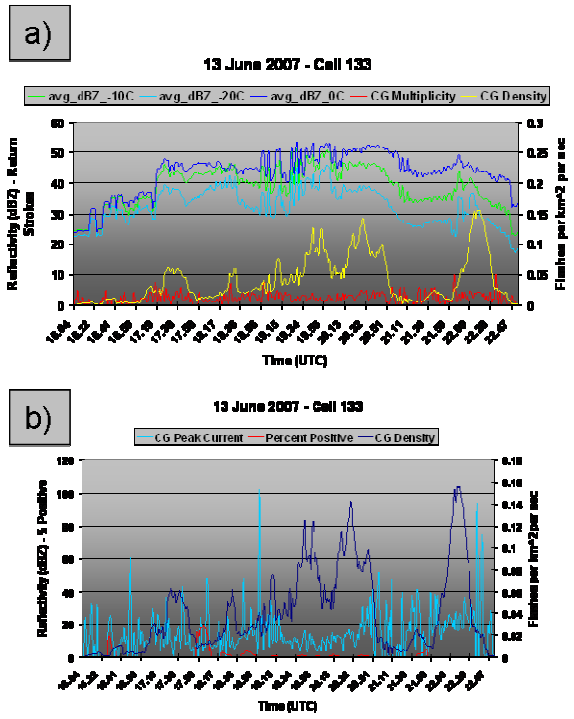

Fig. 7 a) Time series of average dBZ at the melting level (dark blue), at -10 °C (green), and at -20 °C (light blue) as well as average cloud-to-ground multiplicity (red) and flash density (yellow) throughout the life cycle of cell 133 b) Average CG peak current (light blue), flash density (dark blue), and percentage of +CG for the same period.

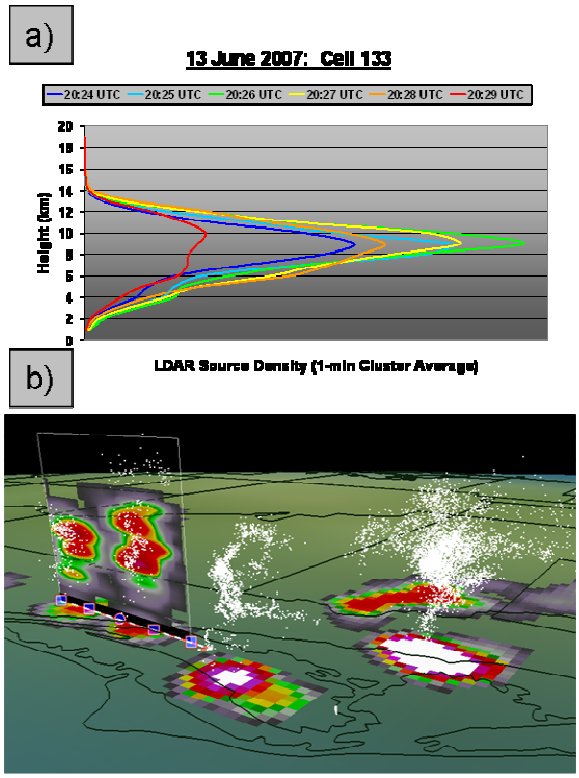

Fig. 8 a) Vertical distribution of LDAR sources at 1 min intervals during a radar volume scan (i.e., density at 1 km intervals between 1 and 20 km) b) Cross-section of LDAR sources (white dots) and the WDSS-II derived density fields as well as plan-view of LDAR source density at 9 km AGL.

grids are computed for 1-, 5-, and 10-min periods, and are reported at 1-min intervals. Figure 7 provides examples of the CG characteristics and radar-derived parameters that we are calculating within WDSS-II. Average reflectivity values at 0 °C, -10 °C, and -20°C are compared with the average CG multiplicity and flash density within an individual storm (Fig. 7a). The average peak current, flash density, and polarity of CG flashes within the same storm are depicted in Fig. 7b. These newly computed fields allow us to take full advantage of the CG characteristics currently reported by the NLDN.

Although CG lightning provides a useful indicator of storm intensity, IC lightning data offer a much more detailed view of storm-scale processes. The few LDAR/LMA networks currently in existence provide localized coverage, so our dataset will include storms from various LDAR/LMA networks to investigate regional influences. The LMA network in Sterling, VA is the focus of our research; however, we also examine storms within the Huntsville, AL and Kennedy Space Center, FL regions.

The original WDSS-II derived IC lightning parameters include the source or flash density at specified levels (e.g., 1-20 km), the Vertically Integrated LMA (VILMA), the Maximum LMA density (MaxLMA), the Height of MaxLMA, and the LMA Layer Average. These parameters describe both the amount and distribution of IC lightning, and allow us to take full advantage of the data's fine temporal resolution (seconds). An important consideration of

our research is the value added by the vertical dimension of lightning data.

Fig. 8a provides an example of the temporal resolution provided by LDAR/LMA data. Each colored line indicates the vertical profile of the LDAR column at 1 min intervals during a single WSR-88D volume scan. A snap shot of the data is shown in Fig. 8b. The individual LDAR sources (white dots) are combined to create LDAR source densities (cross-section) at 1 km intervals from the surface to 20 km AGL.

The GOES-R Global Lightning Mapper (GLM) will provide 2-D total lightning data at a resolution of approximately 10 km (Goodman et al. 2008). We are examining GLM proxy parameters obtained from LDAR/LMA data to provide risk reduction for the GLM. We modify and combine NLDN and LDAR/LMA data within WDSS-II by employing data mining procedures described below. Fuzzy logic techniques currently are being evaluated to develop new GLM proxy parameters. These tools will assign weights to existing lightning variables in order to develop new proxy parameters. The resulting products will exploit both the spatial and temporal (2-D) aspects of total lightning data.

The GLM will provide total lightning data to all NWS WFOs for the first time. Therefore, it is important to determine the suitability of GLM data in native form for assessing storm severity. These data will be used in conjunction with data from existing lightning detection networks and other remote sensing sources to develop new tools for NWS forecasters.

WDSS-II contains an algorithm that uses Hierarchical K-Means clustering to identify and track mesoscale features as small as 10 km^2 (Lakshmanan and Smith 2008). The original National Severe Storms Laboratory (NSSL) algorithm (w2segmotionll) clusters 2-D areas containing average composite reflectivity values greater than 30 dBz. It then computes higher order parameters and employs a user-created decision tree to determine the type of storm (i.e., isolated supercell, line, pulse, or non-severe). These clusters, depicted in Fig. 9, are used to mine data from additional gridded fields (i.e.,

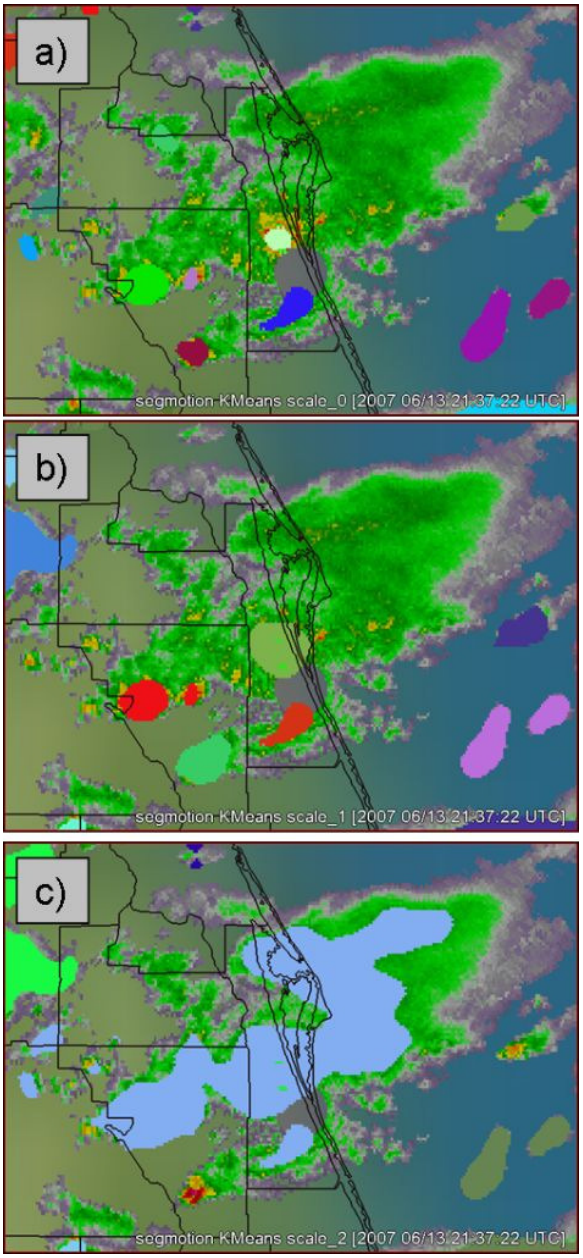

Fig. 9 K-Means clusters for a) Scale 0 b) Scale 1 and c) Scale 2 overlain on composite reflectivity. Scales result from user defined thresholds and are used to track changes in additional gridded fields along the storm's path. Scale 0 identifies the smallest features; scale 2 classifies only the largest or most coherent features; and scale 1 falls in between.

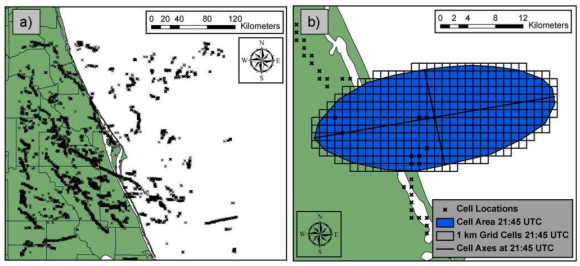

Fig. 10 a) The locations of every cluster during each 1 min period for an 8 h analysis period b) Idealized cluster depicting a grid spacing of 1x1 km. The data mining algorithm (w2segmotionll) tracks the average, count, maximum, minimum, and/or standard deviation of given parameters within the cluster.

LMA and/or NLDN), that are stored in a final storm database. We have modified this algorithm to cluster reflectivity based on various thresholds (e.g., 10 dBZ) and to compute higher order parameters along the storm's path. These higher order fields include the newly developed lightning parameters described previously.

Figure 9 illustrates the "scale" issue when defining storms using the w2segmotionll algorithm. The K-Means clusters are created at three different scales, can be based on any gridded field, and employ user-specified thresholds. The w2segmotionll algorithm allows the user to determine the number of pixels comprising each cluster at each of the scales, as well as the parameter and thresholds to be used. Scale 0 identifies the smallest features (Fig. 9a), scale 2 identifies only the largest or most coherent features (Fig. 9c), and scale 1 (Fig. 9b) falls between. It is very important to fine tune the K-Means options to ensure that individual storms are indeed separate from surrounding features.

Basic grid operations (i.e., count, minimum, maximum, average, and standard deviation) can be applied to any parameter within the K-Means cluster. Fig. 10 depicts an idealized example of how these calculations are made. Each 1×1 km grid cell contains a value for all desired fields; these grid operations then are used to compute the desired parameters within each cluster (e.g., the average or maximum of all 1×1 km grid cells bounded by the cluster). Additionally, the w2segmotionll algorithm allows the analyses of temporal trends (time delta), the computation of lifetime statistics (column operations), and fuzzy logic calculations. Current applications for this algorithm, using the aforementioned lighting and radar fields, are described in the following section.

4. Analysis

Storms from many different geographic regions and atmospheric conditions must be examined to develop robust statistical relationships between lightning and radar parameters and storm intensity. To inspect the large number of storms that is needed, it is important to streamline the development and analysis of our database. Therefore, we are automating the procedures from database creation through the visualization of individual storms. This allows us to minimize manual inspection and to transition away from the more typical case study mode. However, in order to maximize accuracy, individual cases still are examined to physically and visually confirm the results.

Figure 11 illustrates the combination of total lightning and radar data within an individual storm cell. Mean and maximum reflectivities are plotted in Fig. 11a along with the average CG flash density and LDAR source density within a given layer. These same fields are displayed visually in Fig. 11b. The cross-section depicts LDAR source density, while the plan view contains LDAR flash initiation points, CG flash locations, and the CG flash density. The combination of these fields within an individual storm allows statistical relationships to be determined.

Our procedures currently are automated from the retrieval of radar, RUC-derived, and lightning data through the output of individual storms. The final step in this process will be developing storm query and display procedures. We will determine the storm track and compute the average, maximum, and minimum distance from the LDAR/LMA network and the WSR-88D. This will ensure that all data are consistent. The storm query process also will consider storm duration, and allow us to define entire storms in addition to quickly developing features. Based on user-defined attributes within the query program, each final storm folder from

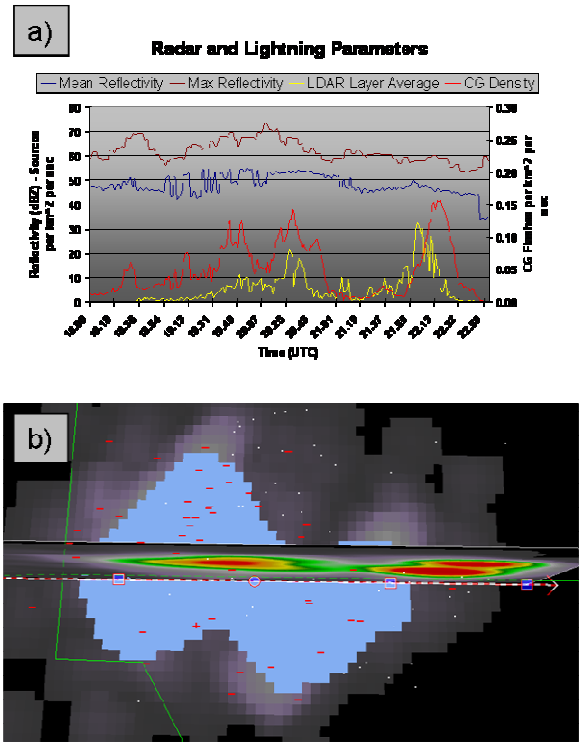

Fig. 11 a) Example of combining radar and lightning parameters computed during the life cycle of a storm (Mean and Maximum reflectivity, LDAR layer average source density, and cloud-to-ground flash density b) View slightly south of nadir, looking down on a cross-section view of LDAR source density and a plan view of LDAR flash initiation points, CG flash locations, and CG flash density.

a given analysis period will contain animations and plots illustrating the most significant storm features and events.

Lang and Rutledge (2008) noted that modularity is key when interrogating large radar and lightning datasets. This is our most desired attribute, and represents the foundation for our research. Regional and seasonal variations in the correlation between lightning and radar require our scheme to be applicable to different geographical regions. Also, our procedure must provide a framework for continuing improvement. This will facilitate the incorporation of new technology (i.e., new data

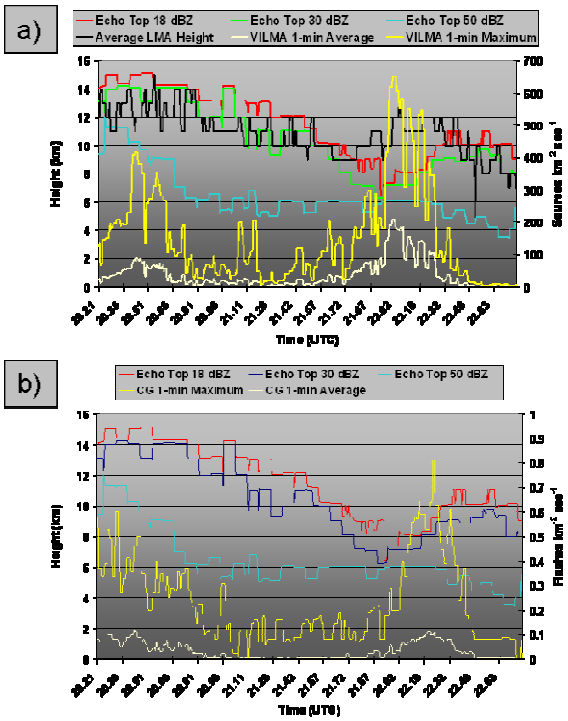

Fig. 12 Example of missing data, cell 133 tracked directly over the radar during peak lightning production a) Echo top at 18 dBZ (red), 30 dBZ (green), and 50 dBZ (blue) – Average Vertically Integrated LMA (VILMA), maximum VILMA, and the average height of maximum LDAR density b) Echo top at 18 dBZ (red), 30 dBZ (dark blue), and 50 dBZ (light blue) – Maximum CG flash density and average CG flash density.

sources) and the inclusion of developing knowledge about storms.

An example of the importance of modularity is depicted in Fig. 12. Cell 133 tracked directly over the radar during peak lightning production. This produced a lowering of radar defined echo tops when the opposite should be expected. During that time, however, the average height of the maximum LDAR/LMA density is observed to shift upward. To account for such occurrences, our guidance products will utilize the currently available parameters to make the most accurate determination of storm severity. Thus, if a data source is missing, we will leverage the remaining sources to assist in the warning decision process.

Naturally, the level of confidence will be affected and must be defined within the guidelines.

As the final aspect in our research, linear regression will be used to derive physical relationships between lightning, radar, and storm severity. Determining the optimum parameters is the first step. Since the combination of possible parameters is virtually infinite, statistical analyses will be used to determine the best relationships and combinations of parameters. The selected parameters will be related to storm type to observe trends within the radar and lightning fields. These relationships then will be used to observe and quantify the 3-D development of many severe and non-severe storms.

Statistical analyses will help reduce the vast number of possible radar, lightning, and RUC parameters, and will be used to develop probabilistic forecasts of severity. The inclusion of lightning data allows us to quantify storm severity at increasing distances from the radar. Quantifying the lightning patterns within severe and non-severe storms will provide forecasters with an additional tool to more accurately determine both the onset and severity of inclement weather. Used with traditional radar parameters, these statistical procedures seek to increase the warning lead time for severe weather events.

5. Conclusions

Utilizing the WDSS-II software, we seek to shift from the case study mode to develop a large database of storms which contains many radar-, lightning-, and RUC-derived parameters. This enhanced database will allow us to derive more robust storm-scale relationships between total lightning, radar, and storm severity. Our focus is on the warning decision support process, and we hope to package the resulting algorithms and guidelines for use at all NWS WFOs.

Our major goal is to help insure that the NWS is using current lightning resources to best advantage when assessing severe weather events. Additionally, the GLM risk reduction component will help ensure that the NWS is prepared to utilize the upcoming total lightning data for determining severe weather potential.

6. Acknowledgements

This work is funded by the National Environmental Satellite, Data, and Information Service (NESDIS), National Oceanic and Atmospheric Administration (NOAA). Specifically, we want to thank Steve Goodman and the GOES-R Global Lightning Mapper (GLM) risk reduction team for their support.

Also, the lead author would like to thank Dr. Valliappa Lakshmanan, aka “Lak”, for the many hours he spent investigating how I managed to break his carefully designed software. Thanks also are owed to Travis Smith, Kurt Hondl, Greg Stumpf, and the other scientists at the National Severe Storms Laboratory (NSSL) and the Cooperative Institute for Mesoscale Meteorological Studies (CIMMS) who have helped to design, maintain, and promote the WDSS-II software.

7. References

Biggar, D.G., 2002: A case study of positive strike dominated supercell thunderstorm that produced an F2 tornado after undergoing a significant cloud-to-ground lightning polarity shift. *Nat. Wea. Digest*.

Carey, L.D., and S.A. Rutledge, 2003: Characteristics of cloud-to-ground lightning in severe and nonsevere storms over the central United States from 1989-1998. *J. Geophys. Res.*, doi:10.1029/2002JD002951.

Carey, L. D., S. A. Rutledge and W. A. Petersen, 2003: The relationship between severe storm reports and cloud-to-ground lightning polarity in the contiguous United States from 1989-98. *Mon. Wea. Rev.*, **131**, 1211-1228.

Carey, L.D., and K.M. Buffalo, 2007: Environmental control of cloud-to-ground lightning polarity in severe storms. *Mon. Wea. Rev.*, **135,** 1327-1353.

Gatlin, P., 2007: Severe weather precursors in the lightning activity of Tennessee Valley thunderstorms. M.Sc. Thesis, Dept. of Atmos.

Sci. University of Alabama in Huntsville (unpublished).

Goodman, S. J., R. Blakeslee, H. Christian, W. Koshak, J. Bailey, J. Hall, E. McCaul, D. Buechler, C. Darden, J. Burks, T. Bradshaw, and P. Gatlin, 2005: The North Alabama Lightning Mapping Array: Recent severe storm observations and future prospects. *Atmos. Res.*, **76**, 423-437.

Goodman, S. J., R. J. Blakeslee, W. Koshak, W. Petersen, D. E. Buechler, P. R. Krehbiel, P. Gatlin, and S. Zubrick, 2008: Pre-launch algorithms and risk reduction in support of the Geostationary Lightning Mapper for GOES-R and beyond. Paper 3.3 *, Third Conference on Meteorological Applications of Lightning Data*, New Orleans, Amer. Meteor. Soc.

Kane, R. J., 1991: Correlating lightning to severe local storms in the northeastern United States. *Wea. Forecasting,* **6,** 3-12.

Knapp, D., 1994: Using cloud-to-ground lightning data to identify tornadic thunderstorm signatures and nowcast severe weather. *Nat. Wea. Dig.*, Vol. 19, No. 2, 35-42.

Lang, T. J. and S. A. Rutledge, 2008: A statistical framework for the analysis of large lightning and radar datasets. Paper 4.8 *, Third Conference on Meteorological Applications of Lightning Data*, New Orleans, Amer. Meteor. Soc.

Lakshmanan, V., T. Smith, G.J. Stumpf, and K. Hondel, 2007: The warning decision support system-integrated information (WDSS-II). *Wea. Forecasting*, 22, 592-608.

Lakshmanan, V. and T. Smith, 2008: Data mining storm attributes from spatial grids. *J. Ocean and Atmos. Tech.*, pp, under review.

MacGorman, D. R., and D. W. Burgess, 1994: Positive cloud-to-ground lightning in tornadic storms and hailstorms. *Mon. Wea. Rev.*, **122**, 1671-1697.

MacGorman, D.R., and C.D. Morgenstern, 1998: Some characteristics of cloud-to-ground lightning in mesoscale convective systems. *J. Geophys. Res.*, 103, 14011-14023.

Montanya, J., S. Soula, and N. Pineda, 2007: A study of the total lightning activity in two hailstorms. *J. Geophys. Res.*, **112**, D13118, doi:10.1029/2006JD007203.

Steiger, S. M., R. E. Orville, and L. D. Carey, 2007: Total lightning signatures of thunderstorm intensity, Part I: Supercells. *Monthly Weather Review*, **135**, 3281-3302**.**

Williams, E. R., B. Boldi, A. Matlin, M. Weber, S. Hodanish, D. Sharp, S. Goodman, R. Raghavan, and D. Buechler, 1999: The behavior of total lightning activity in severe Florida thunderstorms. *Atmos. Res.*, **51**, 245-265.