Lagrangian Analyses of Rainfall Structure and Evolution for Organized Thunderstorm Systems in the Urban Corridor of the Northeastern United States

JUNE K. YEUNG, JAMES A. SMITH, AND MARY LYNN BAECK
Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey

GABRIELE VILLARINI
IIHR–Hydroscience and Engineering, The University of Iowa, Iowa City, Iowa

(Manuscript received 23 May 2014, in final form 18 February 2015)

ABSTRACT

In this study, a climatology of the structure and evolution of rainfall for organized thunderstorm systems in the urban corridor of the northeastern United States is developed. These storm systems are major agents of flash flooding for urban regions of the northeastern United States and, more generally, for the United States east of the Rocky Mountains. The analyses are motivated by problems that center on characterizing flash flood hazards. The authors focus on spatial heterogeneities of rainfall associated with urbanization in a region of complex landscape including mountainous terrain and land–water boundaries along the geometrically complex coastline of the New York City–New Jersey metropolitan region. The sample of storms selected for investigation consists of the 50 days from April to September 2001–09 with the largest cloud-to-ground lightning flash density derived from National Lightning Detection Network (NLDN) observations over the study region. Storm-tracking analyses of 3D radar reflectivity fields are performed for the 50 storm days and used to develop a Lagrangian climatology of storm structure and evolution for the study region. Rainfall analyses for the 50 storm days are based on high-resolution (1 km, 15 min) bias-corrected radar rainfall fields developed from the Hydro-NEXRAD system. The analyses suggest that complex terrain and land–water boundaries have large impacts on Lagrangian storm properties. Areas of increased heavy rainfall and lightning flash density over New York City were identified. The authors found evidence for changing storm structure as thunderstorms pass over New York City, but little evidence that thunderstorms split as they approach New York City.

1. Introduction

In this paper, we develop a regional climatology of rainfall and storm evolution for severe thunderstorm systems in the urban corridor of the northeastern United States. In particular, we present a Lagrangian climatology of storm initiation, storm structure, and storm motion over a region that encompasses the New York City metropolitan area. The region (Fig. 1) is the most densely populated area of the United States, with 27,000 people per square mile in New York City and 1120 people per square mile in New Jersey (U.S. Census Bureau 2010). Our study region is bounded by the Appalachian Mountains to the west, a geometrically complex coastline and the Atlantic Ocean to the east, and a mix of mainly deciduous broadleaf forest, cropland–woodland mosaic, and urban areas throughout the region. The major urban centers of Philadelphia and New York City are connected by a corridor of high-density urban development.

We examine organized thunderstorm systems as dominant agents of flash flooding in urban environments of the United States east of the Rocky Mountains (Smith et al. 2001, 2005; Doswell et al. 1996; Ntelekos et al. 2007, 2008; Wright et al. 2012). The 8 August 2007 storm in New York City (see section 3b), which produced flash floods that closed the subway system and a tornado ranked as 2 on the enhanced Fujita scale (EF2), illustrates the potential of organized thunderstorm systems to disrupt major cities of the eastern United States (Lombardo and Colle 2011). In Ntelekos et al. (2007) and Ntelekos et al.
(2008), it was shown that organized thunderstorm systems are responsible for the majority of urban flash floods in Baltimore, Maryland; that the regional occurrence of flash flooding was linked to the intensity of the thunderstorm systems [represented as a regional cloud-to-ground (CG) lightning flash density]; and that organized thunderstorms producing flash floods in Baltimore exhibited characteristic life cycles involving storm initiation over the eastern margin of the central Appalachians and interaction of storm systems with sea breeze and bay breeze circulations. In a recent study, Smith et al. (2014, manuscript submitted to *J. Hydrometeor.*) use lightning data from the National Lightning Detection Network (NLDN) to show that thunderstorms account for the overwhelming majority of flash floods in urban watersheds for the United States east of the Rocky Mountains. In the New York–New Jersey metropolitan region, thunderstorms account for more than 80% of flood peaks exceeding a unit discharge of $1 \text{m}^3 \text{s}^{-1} \text{km}^{-2}$ [unit discharge is discharge ($\text{m}^3 \text{s}^{-1}$) divided by drainage area ($\text{km}^2$); see Smith et al. (2005) for additional discussion] at U.S. Geological Survey (USGS) stream gauging stations in urban watersheds. This value increases to more than 90% for urban regions in much of the central United States. Based on these previous studies, we focus our analyses on thunderstorm systems that produce large lightning flash densities over the study region. In particular, we will examine the climatology of the 50 largest thunderstorm days during the period from 2001 to 2009.

The role of urbanization in modifying the regional rainfall climatology has received considerable attention, with a focus on urban modification of convective rainfall extending back to the Metropolitan Meteorological Experiment (METROMEX) in St. Louis (see Changnon 1988; Shepherd 2005; Dixon and Mote 2003; Mote et al. 2007; Shem and Shepherd 2009; Miao et al. 2011; Niyogi et al. 2011). Previous studies have examined elements of the warm season rainfall climatology of the New York City metropolitan region. An important theme of recent studies has been the interaction of urbanization impacts with circulation features tied to the complex land–water boundaries and terrain of the region (Colle and Novak 2010; Colle and Yuter 2007; Lombardo and Colle 2010; Wasula et al. 2002; Bornstein and Thompson 1981; Thompson et al. 2007). The interplay of urbanization and complex terrain has also played a central role in recent studies of warm season rainfall in the Baltimore metropolitan region (Smith et al. 2012; Ntelekos et al. 2008) and in the Houston metropolitan region (Carter et al. 2012). For both New York City and Baltimore, analyses of individual storm events have suggested that the impacts of urbanization on the atmospheric boundary layer modify the structure and evolution of thunderstorm systems as they approach and pass over the urban environment.
(Bornstein and LeRoy 1990; Smith et al. 2005; Ntelekos et al. 2008). To address these questions, we will perform Lagrangian analyses of storm structure, motion, and evolution.

Niyogi et al. (2011) developed a Lagrangian climatology of storm properties based on manual tracking and analysis of 91 thunderstorms that passed over the Indianapolis metropolitan region during a 10-yr period (2000–09). Their results show that urbanization can alter the structure, motion, and intensity of thunderstorm systems and illustrate the utility of a Lagrangian perspective in examining urban impacts on rainfall. In this study, we perform Lagrangian analyses of storm properties using the Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN) storm-tracking algorithms (Dixon and Wiener 1993) with 3D volume scan reflectivity observations from the Fort Dix, New Jersey (KDIX), Weather Surveillance Radar-1988 Doppler (WSR-88D). Automated storm-tracking procedures have the advantage of avoiding subjective decisions in storm identification, tracking, and analysis, as well as significantly increasing the ability to build larger climatological datasets of storm properties. Interpretations of automated storm-tracking procedures, however, must be made in light of the algorithm structure and parameters used in implementing the procedures [see Dixon and Wiener (1993) and Javier et al. (2007) for discussion of climatological analyses of storm properties based on the TITAN storm-tracking algorithms].

Rainfall analyses are based on high-resolution (1-km horizontal resolution and 15-min time resolution) rainfall fields that we derive from volume scan KDIX reflectivity observations using the Hydro-NEXRAD rainfall algorithms (see Krajewski et al. 2011). Hydro-NEXRAD rainfall fields have been used for climatological analyses of rainfall in urban regions based on long-term continuous rainfall records in Baltimore (Smith et al. 2012); Atlanta, Georgia (Wright et al. 2012); and Charlotte, North Carolina (Wright et al. 2014b). They have also been used, as in this study, for examining urban impacts on rainfall in Charlotte (Wright et al. 2014a) and Milwaukee, Wisconsin (Yang et al. 2013), based on storm samples. Rainfall estimation for storm samples consisting of organized thunderstorm systems has the advantage of avoiding the most severe problems of radar range degradation associated with stratiform precipitation (Smith et al. 1996; Baeck and Smith 1998). The Fort Dix WSR-88D, unlike some of the surrounding radars, does not suffer from problems associated with blocked radials.

Rainfall variability in time and space has been viewed from a Lagrangian perspective in studies extending back to the point process formulation of LeCam (1961) based on “storm cells” (see also Waymire et al. 1984; Smith and Karr 1985; Smith and Krajewski 1987). A major obstacle to advances in understanding rainfall variability has been the inability to develop statistical procedures for examining rainfall from a Lagrangian perspective (see, e.g., Smith and Karr 1990). In this paper, we develop methods that are broadly applicable for characterization of rainfall variability from volume scan radar reflectivity observations.

The research questions that motivate this study are as follows:

1) What are the characteristic spatial and temporal properties of rainfall for organized thunderstorm systems in the New York–New Jersey region?
2) How do distributional properties of storm initiation, storm motion, and storm splits for organized thunderstorm systems vary over the New York–New Jersey region?
3) How are landscape features, especially those associated with urbanization (but also mountains and land–water boundaries), linked to storm evolution and the distribution of heavy rainfall?

The organization of this paper is as follows. In section 2, we introduce key features of the study area, data, sample of storm days, and the storm-tracking system used to identify and compute storm statistics. In section 3, we present composite analyses of rainfall, lightning, and Lagrangian storm properties based on our sample of 50 storm days. Key conclusions of this study are discussed in section 4.

2. Methodology and study region

The study domain (Fig. 1) is a 51,984 km² area (228 km × 228 km) that encompasses the New York City metropolitan region and includes portions of the states of New York, New Jersey, Pennsylvania, and Connecticut. The study region is contained within the radar umbrella of the Fort Dix WSR-88D radar, which is located in Monmouth, New Jersey. Major terrain features in the study area include the Appalachian Mountains to the west, a geometrically complex coastline and the Atlantic Ocean to the east, and the major urban areas of New York City and Philadelphia.

We use CG lightning data from the NLDN (see Orville and Huffines 2001; Villarini and Smith 2013) as the basis for developing our sample of storm days. As noted in section 1, our sample of storm days reflects events that are major agents of flash flooding. We use NLDN observations during the period from 2001 to 2009 for the “warm season” months of April–September. We ranked each day (1200–1200 UTC) by the CG lightning...
flash density over the study region (Fig. 2) and selected the top 50 days as our sample for storm analyses (Table 1). Several storm days had missing radar observations, and the next storm days in the ranking were used to make up the 50-storm sample. We have carried out analyses to assess sensitivity of results to details of the domain configuration and sample size. These analyses suggest that the conclusions are not overly sensitive to changes in domain size and sample size. Future studies should, however, more closely examine the dependence of climatological analyses of thunderstorm properties on spatial scale and questions centering on whether extreme events are different from more common events.

Different metrics have been used to define intense convection, including lightning flash density (e.g., Murray and Colle 2011), reflectivity extremes in volume scan radar observations (Smith et al. 2001), satellite observations from microwave imager and precipitation radar (Nesbitt and Zipser 2003), and severe weather reports (see Wasula et al. 2002; Lombardo and Colle 2010). The use of CG lightning in this study avoids problems associated with radar artifacts (e.g., potential radar beam blockage, ground clutter, and uncertainties associated with radar bias-correction factor computations), density of the rain-gauge network, and population bias associated with severe storm reports [see Wasula et al. (2002) for discussion]. The correlation between lightning flash density and rainfall is well documented [see, e.g., Tapia et al. (1998), and references therein], as is the link between lightning flash density and flash flooding [see Ntelekos et al. (2007), and references therein].

We derived radar rainfall estimates for this study at 15-min time resolution and 1-km spatial resolution using

![Fig. 2. The number of CG lightning strikes observed within the study domain (see Fig. 1) for each of the top 50 convective storm days (1200–1200 UTC) selected for this study. Dates that correspond to each storm ranking are given in Table 1.](image_url)

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Date</th>
<th>CG lightning strikes</th>
<th>Ranking</th>
<th>Date</th>
<th>CG lightning strikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27 Jul 2008</td>
<td>23 621</td>
<td>26</td>
<td>7 Aug 2007</td>
<td>7729</td>
</tr>
<tr>
<td>2</td>
<td>26 Jun 2009</td>
<td>23 489</td>
<td>27</td>
<td>1 Jul 2005</td>
<td>7711</td>
</tr>
<tr>
<td>3</td>
<td>1 Jun 2006</td>
<td>20 758</td>
<td>28</td>
<td>2 Aug 2008</td>
<td>7570</td>
</tr>
<tr>
<td>4</td>
<td>23 Jul 2008</td>
<td>19 286</td>
<td>29</td>
<td>1 Jul 2001</td>
<td>7401</td>
</tr>
<tr>
<td>5</td>
<td>14 Aug 2005</td>
<td>18 316</td>
<td>30</td>
<td>12 Jul 2006</td>
<td>7063</td>
</tr>
<tr>
<td>6</td>
<td>29 Jul 2009</td>
<td>16 315</td>
<td>31</td>
<td>11 Aug 2008</td>
<td>7020</td>
</tr>
<tr>
<td>7</td>
<td>10 Aug 2001</td>
<td>14 126</td>
<td>32</td>
<td>2 Jul 2006</td>
<td>6897</td>
</tr>
<tr>
<td>8</td>
<td>11 Aug 2004</td>
<td>13 884</td>
<td>33</td>
<td>10 Aug 2008</td>
<td>6728</td>
</tr>
<tr>
<td>9</td>
<td>14 Jun 2008</td>
<td>13 219</td>
<td>34</td>
<td>8 Jun 2008</td>
<td>6558</td>
</tr>
<tr>
<td>10</td>
<td>12 May 2004</td>
<td>11 659</td>
<td>35</td>
<td>11 Jul 2009</td>
<td>6546</td>
</tr>
<tr>
<td>11</td>
<td>21 Jul 2006</td>
<td>10 856</td>
<td>36</td>
<td>21 Jul 2003</td>
<td>6433</td>
</tr>
<tr>
<td>12</td>
<td>6 Jun 2005</td>
<td>10 782</td>
<td>37</td>
<td>29 Jun 2005</td>
<td>6120</td>
</tr>
<tr>
<td>13</td>
<td>10 Jun 2008</td>
<td>10 780</td>
<td>38</td>
<td>14 Aug 2008</td>
<td>5973</td>
</tr>
<tr>
<td>15</td>
<td>22 Jul 2003</td>
<td>10 059</td>
<td>40</td>
<td>14 Jul 2004</td>
<td>5805</td>
</tr>
<tr>
<td>16</td>
<td>17 Aug 2007</td>
<td>9950</td>
<td>41</td>
<td>8 Jun 2009</td>
<td>5735</td>
</tr>
<tr>
<td>17</td>
<td>18 Jul 2006</td>
<td>9715</td>
<td>42</td>
<td>25 Aug 2006</td>
<td>5584</td>
</tr>
<tr>
<td>18</td>
<td>29 Jul 2007</td>
<td>9676</td>
<td>43</td>
<td>1 Jul 2009</td>
<td>5479</td>
</tr>
<tr>
<td>19</td>
<td>10 May 2004</td>
<td>9062</td>
<td>44</td>
<td>18 Jul 2005</td>
<td>5288</td>
</tr>
<tr>
<td>20</td>
<td>9 Sep 2008</td>
<td>8568</td>
<td>45</td>
<td>5 Aug 2005</td>
<td>5164</td>
</tr>
<tr>
<td>21</td>
<td>3 Aug 2007</td>
<td>8530</td>
<td>46</td>
<td>9 Jun 2009</td>
<td>5065</td>
</tr>
<tr>
<td>22</td>
<td>27 Jul 2006</td>
<td>8270</td>
<td>47</td>
<td>30 Aug 2003</td>
<td>5030</td>
</tr>
<tr>
<td>23</td>
<td>27 Jul 2005</td>
<td>7976</td>
<td>48</td>
<td>27 Jul 2007</td>
<td>4936</td>
</tr>
<tr>
<td>24</td>
<td>26 Jul 2009</td>
<td>7949</td>
<td>49</td>
<td>26 Jul 2008</td>
<td>4724</td>
</tr>
<tr>
<td>25</td>
<td>2 Aug 2009</td>
<td>7765</td>
<td>50</td>
<td>7 Aug 2006</td>
<td>4723</td>
</tr>
</tbody>
</table>
volume scan reflectivity observations from the Fort Dix WSR-88D radar using the Hydro-NEXRAD system (Smith et al. 2012; Krajewski et al. 2008; Seo et al. 2011; Smith and Krajewski 1991). Multiplicative bias correction of radar rainfall fields was applied for each of the 50 days by taking the ratio of the mean rainfall from rain gauge observations over the entire study area to the collocated radar observations. For each gauge, only hours with nonzero rainfall for both gauge and collocated radar observations were included in the computation. In this study, rain gauge observations maintained by the Office of the New Jersey State Climatologist were used to compute the daily multiplicative bias. The network provides good coverage of the region for the purposes of multiplicative bias correction [see Smith et al. (2012) for related analyses and discussion]. We use rainfall analyses based on a single radar because of the mature algorithms for single radar estimation, especially using the Hydro-NEXRAD algorithms, and because of blockage problems associated with surrounding radars. Problems associated with range-dependent errors in rainfall estimates are much less severe for organized thunderstorm systems than for stratiform precipitation (Baek and Smith 1998).

The TITAN storm-tracking algorithms were used to identify and track storm elements from 3D reflectivity fields (Dixon and Wiener 1993). A storm element is defined as a contiguous volume of at least 50 km$^3$ with radar reflectivity exceeding 45 dBZ. We used a 45 dBZ threshold (as opposed to lower thresholds) to obtain convective storm elements and minimize the impact of brightband contamination from stratiform precipitation. Storm tracking is based on a cross-correlation algorithm that compares 3D reflectivity fields (Dixon and Wiener 1993) from sequential radar volume scans. In addition to tracking storm elements from one volume scan time to the next, the TITAN storm-tracking algorithm identifies splits and mergers of storm elements. Splits occur when a storm element produces two storm elements (children) in the following volume scan. Mergers occur when a storm element is the product of two storm elements (parent) from the preceding volume scan. TITAN analyses of storm properties were used in Javier et al. (2007) to develop a climatology of thunderstorm properties in the Front Range of the Rocky Mountains. In Tapia et al. (1998), storm-tracking analyses using TITAN are coupled with analyses of cloud-to-ground lightning observations to examine the climatology of severe thunderstorms in Florida. These and other studies with similar automated storm-tracking algorithms (see, e.g., Peleg and Morin 2012) illustrate and discuss the sensitivity of results to algorithm structure and parameters.

We use storm-tracking analyses to examine splits, mergers, initiation, and decay locations of storm elements. A “true” initiation occurs when a storm element reaches the defined volume and reflectivity thresholds for a storm cell and the storm element does not have any parent cells. Analyses of “storm initiation” using the TITAN algorithm provide the time and location when the storm cell first exceeds the reflectivity and volume thresholds; this formulation of storm initiation provides a later time and larger size storm element than definitions based on first echo definitions [see Wilson and Mueller (1993) for discussion of storm initiation and reflectivity thresholds]. A true storm decay occurs when a storm element does not have any children cells (i.e., no new cells form as a result of the “decay” of this element). Since some storms initiate as a result of splits and mergers, the number of storm initiations does not necessarily equal the number of decay events. Storm duration is defined as the time elapsed between true initiation and the last true decay of this track. A split occurs when a storm element has two or more children cells, and a merger has occurred when a cell has two or more parent cells.

Tracked storm elements in TITAN are identified by “simple” and “complex” track numbers. Simple storm tracks have no splits or mergers, and complex tracks include all simple tracks that are parents or children of the same cluster of storm tracks during their lifetime. In other words, a number of simple tracks make up a complex track, and all the simple tracks in a complex track are related through splits and mergers. When we describe a major storm element passing over a certain part of our study area, we are referring to a particular complex track and all the associated simple tracks.

3. Results and discussion

a. Overview of the 50 storm days

The 50 storm days (Table 1 and Fig. 2) have CG lightning flash densities that range from a total of 23,621 CG lightning strikes (for a mean flash density of 0.45 CG strikes km$^{-2}$ over the 51,984 km$^2$ study region) for the record storm day of 27–28 July 2008 to 4723 CG strikes (0.09 CG strikes km$^{-2}$) on 7–8 August 2006. The largest 10 days (Fig. 2) account for 37% of all CG strikes. July and August account for 76% of the storm days, with 42% in July and 34% in August (Fig. 3). They are followed by June (18%), May (4%), and September (2%).

There is large interannual variability in the number of days ranking in the top 50 lightning events (Fig. 4), with 10 days in 2008 and none in 2002. This is consistent with previous studies, which also show that interannual
variability is an important element of the climatology of warm season thunderstorm systems (see, e.g., Murray and Colle 2011). The smallest storm counts occurred during the early part of the record. The years 2001, 2003, and 2004 follow 2002 as the years with smallest event counts. Because of the small sample size, it is not possible to assess temporal changes in severe storm occurrences [see Villarini and Smith (2013) for additional discussion].

Storm total rainfall and CG flash density fields for the 50 storms exhibit large spatial variability and reflect spatial heterogeneities in storm properties associated with mountainous terrain, land–water boundaries, and urban land cover (as discussed in the following section). We illustrate storm properties and methods through analyses of two storm events. The 14–15 August 2005 storm (Figs. 5–7) ranked fifth in CG lightning strikes (see Table 1) and produced significant flash flooding over the New York City metropolitan region. The 7–8 August 2007 storm (Figs. 8–10) produced the twenty-sixth largest CG flash density, the most intense tornado (EF2) in New York City on record, and flash flooding that closed the New York subway system.

Storm total rainfall and cloud-to-ground flash density for the 14–15 August 2005 storm (Fig. 5) show local maxima over the New York City region and northeast of the New York City region (compare with composite rainfall and lightning for the 50 storms in the following subsection). The rainfall maximum over New York City of 120 mm is paired with a 220-mm maximum northeast of the city. The CG flash density maximum over the New York City region of 3 CG strikes km$^{-2}$ is displaced to the northwest, relative to the rainfall maximum. The maximum CG flash density of 4 CG strikes km$^{-2}$ is collocated with the rainfall maximum northeast of the city (Fig. 5). In Fig. 5 (and in subsequent contour maps), rainfall and lightning maps are created from 1-km grid cell values of rainfall and lightning flash density using the ArcGIS contour mapping algorithm with a grid smoothing routine.

Time series of storm properties for the dominant complex storm track that passed over New York City for the 14–15 August 2005 storm (Fig. 6) illustrate storm structure, motion, and evolution of an organized thunderstorm system. We show time series of storm area (km$^2$), maximum reflectivity (dBZ), and storm speed (km h$^{-1}$) for the largest simple track (at each time step) within the complex track. We also show time series of the number of simple tracks within the complex track. In Fig. 7, we illustrate storm structure during the 3-h period.
Fig. 5. Storm total (1200–1200 UTC) (top) rainfall accumulation (mm) and (bottom) CG lightning flash density (strikes km$^{-2}$) for the 14–15 Aug 2005 storm. The base map for the figure illustrates topography of the study region, with lower-elevation areas black and the highest areas in white. Urban regions are overlaid in gray.
preceding heavy rainfall over New York City through reflectivity maps at 2.5-km elevation above ground level. The storm experienced a sequence of splits and mergers over its life cycle, with a peak in the number of simple tracks between 0000 and 0130 UTC. As the storm elements approached New York City, multiple mergers occurred, and ultimately a large linear feature of intense convection (a single simple track) passed over the city between 0200 and 0300 UTC (Fig. 7). During this same time period, a second large, linear simple track moved southwest of New York City. Storm area of the largest simple track increased from less than 200 km$^2$ at 0000 UTC 15 August to more than 600 km$^2$ by 0200 UTC 15 August. Maximum reflectivity exceeded 60 dB$Z$ for approximately 5 h (from 2200 UTC 14 August to 0300 UTC 15 August), reaching a maximum of 70 dB$Z$ (indicating large hail) at 0200 UTC 15 August. Storm speed fluctuated between 40 and 60 km h$^{-1}$ (13–17 m s$^{-1}$) from 0000 through 0200 UTC 15 August, with generally decreasing storm speed after 0100 UTC. Storm direction was principally from west to east. There was not a significant increase in the maximum reflectivity or in echo top height (not shown) during the time period when the storm passed over the city. There was no evidence of splitting of storm elements as they approached the city. Storm speed increased rapidly once the storm passed the land–water boundary into the marine environment of the Atlantic Ocean.

Storm total rainfall and cloud-to-ground flash density for the 7–8 August 2007 storm reflect storm structure, motion, and evolution of the storm system (Fig. 8). The spatial maximum in CG flash density of 4.5 CG strikes km$^{-2}$ was displaced to the southeast relative to the rainfall maximum of 90 mm located on the border of northwestern New Jersey and Pennsylvania. Both rainfall and lightning decreased toward the east, and the storm total rainfall was concentrated in northern New Jersey and over New York City.

Time series of storm area (largest simple track), storm speed, and maximum reflectivity for the dominant complex storm track that passed over New York City for the 7–8 August 2007 storm (Fig. 9) show similarities to the 14–15 August 2005 storm, as well as some notable contrasts. In Fig. 10, we illustrate storm structure during the 3-h period preceding heavy rainfall over New York City through reflectivity fields at 2.5 km above ground.
level. This storm also experienced a sequence of splits and mergers over its life cycle. Like the August 2005 storm, the number of simple tracks decreased as the storm approached the New York City metropolitan area, with a consequent increase in the area of the largest simple track. The number of simple tracks peaked at approximately 0700 UTC and decreased systematically because of storm mergers, with a single storm element (simple track) passing over the New York City metropolitan region from 1000 to 1200 UTC (as illustrated in Fig. 10). One distinctive feature of this complex track is the sustained intensity over a long duration. The storm elements that comprised the 7–8 August 2007 storm had simple tracks with maximum reflectivity values greater than 60 dBZ for 10 h. There were not significant changes in the maximum reflectivity, echo top height (not shown), and speed as the storm passed over New York City between 1000 and 1200 UTC.

The August 2005 and August 2007 storms exhibit spatial patterns of rainfall and lightning distribution that reflect systematic spatial heterogeneities associated with organized thunderstorms in the study region (a central topic of the next subsection). Increasing area of the largest simple track of a storm element (i.e., complex track) reflects upscale growth of storm systems over the diurnal life cycle of the storm system.

b. Composite rainfall and lightning analyses for the 50 storm days

The mean lightning flash density for the sample of 50 storm days (Fig. 11) shows local maxima on the southwest boundary of New York City, north of the Philadelphia metropolitan region, and over mountainous terrain along the northwestern boundary of the study region. There is a local maximum in CG flash density, which is north of New York City and west of the New York–Connecticut boundary.

The spatial field of mean rainfall accumulation from the top 50 convective days (Fig. 12) shows local maxima northeast of New York City (closed contour of 22 mm day$^{-1}$ that extends across the Connecticut–New York boundary), New York City (20 mm day$^{-1}$), and
the high-elevation region in the northwestern part of the study domain (20 mm day$^{-1}$). The mean rainfall over the study region ranges from 6 to 22 mm day$^{-1}$, with the largest rainfall accumulations over the northern portion of the domain and minimum rainfall over the Atlantic Ocean and adjacent regions of southern New Jersey. The rainfall maximum north of New York City (Fig. 12) is approximately collocated with the local

---

**Fig. 8.** As in Fig. 5, but for the 7–8 Aug 2007 storm.
maximum in CG flash density (Fig. 11). There is a spatial offset between the maximum rainfall and lightning over New York City, with the maximum in rainfall displaced to the northeast, that is, downwind, of the maxima in lightning. The CG flash density in the mountainous northwestern portion of the study region is displaced south and west of the rainfall maximum and is located farther “downslope” from the maximum terrain elevations than the local maximum in rainfall. Gradients in lightning flash density over Long Island, unlike those for rainfall, are not oriented along the main axis of Long Island.

The New York City rainfall maximum is located at the western margin of Long Island. There is a sharp decrease in rainfall over the Atlantic Ocean that parallels the southern coast of Long Island. Mean rainfall decreases from 14 mm day$^{-1}$ along the southern boundary of Long Island to 10 mm day$^{-1}$ over a distance of approximately 40 km. A broad region of low rainfall accumulations (less than 10 mm day$^{-1}$) with weak gradients transitions to sharply increased rainfall accumulations with large gradients over the New York City metropolitan region. There is a weak minimum in rainfall over the Long Island Sound (north of Long Island) between the broad Long Island maximum and the regional maximum of 22 mm northeast of New York City. Larger gradients in rainfall west of the New York City metropolitan region are collocated with large gradients in terrain in the high-elevation area of the study domain.

The distribution of the heaviest rainfall for the 50 sample days, represented by the probability that rainfall exceeds 25 mm day$^{-1}$ (Fig. 13), exhibits a similar spatial pattern as the mean rainfall. Local maxima in heavy rainfall frequency are located northeast of the New York City metropolitan region and in the New York City region at the western end of Long Island. A broad region of increased heavy rainfall frequency is located in the mountainous portion (northwestern) of the study region. There is a region of low frequency of heavy rainfall over southern New Jersey and over the Atlantic Ocean. The area of minimum frequency of heavy rainfall is along the southern New Jersey coast. Gradients in heavy rainfall frequency from the New York metropolitan region, including Long Island, to the southern portion of the study region are even sharper than the gradients in mean rainfall for the 50 storm days.

The distributions of the 24-h rainfall accumulation across the study domain for the 50 events are summarized through box plots in Fig. 14. In Fig. 14, the 25th percentile, the median, and the 75th percentile rainfall accumulations are shown for each storm, with the bottom of each box representing the 25th percentile of
rainfall over all grid points in the domain, the open circle representing the median value, and the top of the box representing the 75th percentile. The daily rainfall accumulations were computed from the bias-corrected Hydro-NEXRAD rainfall fields at 1-km² grid size and 15-min temporal resolution. The median 24-h rainfall accumulation for the 50 events, indicated by the black line across the entire figure, is 8.4 mm. There is large variability in rainfall distribution over the study region. For 12 of the 50 storms, at least 25% of the study region had no rainfall for the day. Events 7, 8, 19, 31, and 49 had generally small rainfall accumulations over the entire domain with 75% of the domain having rainfall accumulations less than the median value of 8.4 mm. There is little relationship between maximum flash density and rainfall distribution. The two largest rainfall days rank thirty-fourth and fiftieth in CG flash density. Five out of the 10 largest days in terms of CG flash density have more than 25% of the area with rainfall accumulations exceeding 20 mm, compared with 23 for the full sample of 50 days.

There is a pronounced diurnal cycle in both CG lightning flash density and rainfall properties for the 50 storm days (Fig. 15). Box plots of spatially averaged rainfall and lightning by hour of the day (Fig. 15) are used to illustrate the diurnal cycle of rainfall and lightning for days with intense convective rainfall [compare with regional analyses presented in Murray and Colle (2011) and national analyses presented in Wallace (1975), Winkler et al. (1988), and Baeck and Smith (1995)]. The mean rainfall and lightning diurnal cycles peak at approximately 2000–2200 local time (0000–0200 UTC). Rainfall was stratified into heavier rain events and all other rain events in the form of the percent of the study domain that exceeds 25 and 1 mm h⁻¹, respectively. This figure shows that the diurnal cycle of 25 mm h⁻¹ tends to peak earlier (at approximately 1700 local time, or 2100 UTC) than the mean rainfall. Periods of intense rainfall from organized thunderstorm systems over the study region are typically followed by extended periods of stratiform precipitation.
Composite analyses for the 50 storm samples reflect some features of previous studies in the region. The broad features of lightning flash density for the continuous warm season climatology of Murray and Colle (2011) are reflected in the composite analyses for the most severe 50 storms (Fig. 11). There are also differences in magnitudes and gradients of rainfall that arise in comparing a continuous climatology with a storm...

![Fig. 11. Mean CG lightning flash density (strikes km$^{-2}$ day$^{-1}$) for the 50 convective days over the study domain. The base map illustrates topography and is overlaid by a grayscale representation of major urban areas.](image1)

![Fig. 12. Mean rainfall (mm) for the 50 convective storm days over the study domain. The base map illustrates topography and is overlaid by a grayscale representation of major urban areas.](image2)
sample climatology. In Smith et al. (2012), it is shown that there are sharper gradients in rainfall for storm sample composites associated with heavy rainfall and flash flooding in the Baltimore metropolitan region than appear in climatological analyses that utilize continuous observations. For rainfall analyses in particular, there are advantages in restricting analyses to organized thunderstorms systems that produce heavy rainfall; some of the error sources that affect continuous radar rainfall records (especially those due to bright band, mixed phase precipitation, and ice phase precipitation) are less of a problem for organized thunderstorm systems (Baeck and Smith 1998).

The influence of circulations linked to land–water boundaries and urban surfaces clearly play an important role in heavy rainfall from organized thunderstorm systems over the New York City metropolitan region (see, e.g., Bornstein and Thompson 1981; Pullen et al. 2007; Holt and Pullen 2007; Colle and Novak 2010). Building on analyses in Colle and Novak (2010), Ryu et al. (2015) show that there are striking diurnal cycles of water vapor flux in the atmospheric boundary layer due to land–water gradients over the mid-Atlantic regions and that the properties of water vapor flux vary markedly for heavy rainfall settings. A combination of novel observational analyses (as in Colle and Novak 2010; Ryu et al. 2015) and numerical experiments (as in Pullen et al. 2007) are needed to attribute physical processes to key features of rainfall and lightning climatology that are identified from the 50-storm composite.
c. Storm-tracking analyses

In this section, we present storm-tracking analyses for the 50 storms. Analyses center on composite analyses, based on all “simple tracks” (see section 2) for the 50 storms. We examine the dominant speed and direction of storm motion, preferred times and locations of storm initiation, decay, splitting, and merging, and we also quantify the statistical distributions of key storm properties, including storm speed, area, volume, duration, and maximum reflectivity.

Storm motion for the 50 storms is dominated by southwest-to-northeast movement of storm elements (Fig. 16). To produce this aggregate map, we identified each storm element (simple track) from all 50 storms that pass through each 1-km grid box in the study region. The mean speed and direction for the grid box is computed by averaging motion vectors for all storm elements in the grid box. Storm cells move faster over the maritime environment of the Atlantic Ocean than over land. Rapid storm motion is one element of low rainfall accumulation over the Atlantic Ocean (see Fig. 12).

The spatial field of storm initiation for the 50 storm days (Fig. 17, top) shows local maxima in New York City, southeastern New York State, west of the Chesapeake Bay in Maryland, and, with a lower frequency, the New Jersey coast, Long Island, and Philadelphia. The maxima in storm initiation are largely due to initiations that are not “true” initiations as defined in section 2 but are associated with preexisting storm elements that cross the defined threshold for volume and reflectivity.

The spatial field of storm splits (Fig. 17, bottom) also shows that the highest frequency of splits is approximately collocated with the areas with the highest frequency of convective activity. The low-lying area between the Appalachian Mountains and the Atlantic coastline experienced the highest frequency of storm activity from the 50 days, which is consistent with previous findings (Murray and Colle 2011). We do not find evidence for increased frequency of storm splits around

![Hourly Rainfall Accumulation (mm)](image)

![% of Domain > 25 mm](image)

![% of Domain > 1 mm](image)

![# of Lightning Strikes](image)

![Fig. 15. Rainfall and lightning diurnal cycles, represented through 24-hourly box plots: (from top to bottom) mean hourly rainfall accumulation over the study region, percent of the study region receiving rainfall accumulation >25 mm, percent of the study region receiving rainfall accumulation >1 mm, and number of lightning strikes over the study region. The circles with a dot in the center represent median values, the thick black boxes represent the 0.75 and 0.25 percentiles, and the thin black lines, the 0.9 and 0.1 ones.](image)
New York City (see also Bornstein and LeRoy 1990). The sample of storms used in this study is dominated by events with strong dynamic forcing and the structure of convection (see, e.g., Fig. 7) is different from the storms examined in Bornstein and LeRoy (1990). Although storm splitting is not prominent around New York City in the TITAN analyses, there is increased frequency of storm initiation around the city, potentially reflecting interactions of storm systems with urban canopy and urban heat island circulations.

The distributions of storm speed for the 50 storm days are shown in box plots in Fig. 18 (for each storm day, more than 100 simple tracks are available for computing box plots). The median value of storm speed over all storm elements from the 50 storms, shown as the horizontal black line in the figure, is 32 km h\(^{-1}\) (8.9 m s\(^{-1}\)). The median value of storm speed, for an individual storm day, ranges from a minimum value of 15 km h\(^{-1}\) to a maximum value of 50 km h\(^{-1}\) (4.2–13.9 m s\(^{-1}\)). The spread in the distribution of storm speeds among storm elements, represented by the interquartile range, can exceed 30 km h\(^{-1}\) (8.3 m s\(^{-1}\)) and is generally greater than 15 km h\(^{-1}\) (4.2 m s\(^{-1}\)). There is not a strong relationship between storm speed and the distribution of storm total rainfall (Fig. 14). The days with largest rainfall accumulation (days 34 and 50) are days with relatively large storm speeds. This suggests that rapidly moving storm elements that repeatedly track over the same region produce the largest point rainfall accumulations in the 50 storm sample, as opposed to slow-moving storm elements (see Schumacher and Johnson 2006).

The distributions of maximum reflectivity values for the 50 storms are shown in box plots in Fig. 19. The median value of maximum reflectivity for all storms, shown as the horizontal black line in the figure, is 56 dBZ. The median value of maximum reflectivity exceeds 53 dBZ (a commonly used threshold for distinguishing the presence of hail) for all 50 storms. Again, there is not a strong link between convective intensity, as...
FIG. 17. Frequency of storm cell (top) initiation and (bottom) splitting for the top 50 convective days. Elevation contours are shown in black.

Density (#/sq. km)

<table>
<thead>
<tr>
<th>Color</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>1.4</td>
</tr>
<tr>
<td>Low</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Fig. 17. Frequency of storm cell (top) initiation and (bottom) splitting for the top 50 convective days. Elevation contours are shown in black.
reflected in maximum reflectivity, and maximum rainfall accumulation. Convective intensity is relatively large for one of the largest rain days (34) and relatively small for another (50). Cotton and Anthes (1992) suggest that low precipitation efficiency of convectively intense thunderstorms reduces their potential for producing heavy rainfall and flash flooding (combined with higher likelihood of rapid storm motion). Smith et al. (2001), on the other hand, argue that, despite relatively low efficiency and rapid motion, supercell thunderstorms represent major flash flood hazards of the United States east of the Rocky Mountains, especially in urban regions, because of their potential for extreme, short-term rainfall rates. The most extreme consequences of flash flooding in the New York City metropolitan area were produced by the 8 August 2007 tornadic thunderstorm. Future studies are needed to assess the links between convective intensity and flash flood potential in urban areas from organized thunderstorm systems.

4. Summary and conclusions

In this study, we examined the climatology of the 50 most intense thunderstorm days (1200–1200 UTC) for the New York City metropolitan region during the period from April to September 2001–09 using high-resolution radar rainfall observations and the TITAN storm-tracking algorithms to objectively characterize statistical properties of storm structure and motion. The 50 days were selected based on the number of lightning strikes observed in the study area on each day in the study period. The storm sample reflects events that are the greatest hazards for flash flooding in the urban corridor of the northeastern United States. We used storm properties computed from the TITAN storm-tracking algorithms in conjunction with spatial fields of radar rainfall observations to characterize the rainfall climatology of organized thunderstorm systems. Key conclusions of this study are as follows.

First, there is pronounced spatial heterogeneity in rainfall for the 50 storm days. Rainfall maxima in composite rainfall analyses for the 50 storm days are located over New York City, north of New York City along the New York–Connecticut border, and in the high-elevation region in the northwestern part of the study domain. Lightning maxima are similar, but with a slight offset, especially over New York City, as the lightning maximum occurs upwind of the rainfall maximum. There is a pronounced seasonal cycle in the occurrence
of storm days, with July and August accounting for more than 75% of the 50 events. There is also a pronounced diurnal cycle in rainfall and lightning flash density for the 50 storm days. Over the diurnal cycle, lightning and heavy rainfall precede mean rainfall and fractional coverage of positive rainfall. Our analyses point to the potential for urban modification of rainfall to play a significant role in determining the spatial distribution of flash flood hazards in urban environments.

Second, Lagrangian analyses of storm motion based on the TITAN storm-tracking analyses for the 50 storm days show that storm elements have a pronounced southwest-to-northeast direction dominating storm motion. The median values of storm area, storm speed, and maximum reflectivity are 25.8 km$^2$, 32 km h$^{-1}$ (8.9 m s$^{-1}$), and 56 dBZ, respectively. Spatial analyses of storm motion show that there is a large increase in storm speed over open ocean. There is large variability in daily rainfall over the study region for the 50 storm days. There is not a strong relationship between speed of storm elements and rainfall, nor is there a strong relationship between convective intensity, as represented by maximum reflectivity of the storm, and total rainfall.

Third, there is little evidence from the TITAN storm-tracking analyses of the 50 storm days that storm elements split when approaching New York City. There is, however, an increased frequency of storm initiation associated with preexisting storm elements. There is relatively low storm frequency over open ocean. The sample of storms examined in this study is typically associated with strong dynamic forcing. Future studies should combine Lagrangian analyses of storm properties with numerical experiments to assess the role of urban heat island circulations and urban canopy effects on organized thunderstorm systems as they pass over urban environments.

Lagrangian analyses of storm structure and evolution, based on manual procedures (Niyogi et al. 2011), have provided useful insights into urban modification of rainfall. In this study, we illustrate the utility of automated storm-tracking procedures for Lagrangian analyses of storm structure, motion, and evolution over urban environments. More generally, Lagrangian analyses of storm properties hold significant potential for examining the temporal and spatial variability of rainfall along the lines envisioned by Le Cam (1961).

Acknowledgments. This work was partially funded by the NOAA Cooperative Institute for Climate Sciences (Grant NOAA CICS NA08OAR4320752), NASA (Grants NNX13AG94G and NNX10AI46G), and the National Science Foundation (Grants CBET-1058027 and EEC-0540832). NLDN data were provided by the NASA Lightning Imaging Sensor (LIS) instrument team and the LIS data center via the Global Hydrology Resource Center (GHRC) located at the Global Hydrology and Climate Center (GHCC), Huntsville, Alabama, through a license agreement with Global Atmospherics, Inc. (GAI). The data available from the GHRC are restricted to LIS science team collaborators and to NASA EOS and TRMM investigators.

REFERENCES


Holt, T., and J. Pullen, 2007: Urban canopy modeling of the New York City metropolitan area: A comparison and validation of


