An Object-Oriented Characterization of Extreme Precipitation-Producing Convective Systems in the Midwestern United States

NATHAN M. HITCHENS,* MICHAEL E. BALDWIN, AND ROBERT J. TRAPP

Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana

(Manuscript received 29 June 2011, in final form 26 September 2011)

1. Introduction

Each year inland flash flooding in the United States poses a significant threat to life, property, and agriculture. Such flooding is often attributed to “extreme precipitation,” which herein will refer to rainfall occurring in 1 h, in amounts exceeding a statistically based threshold. For example, Hitchens et al. (2010) examined rain gauge data in the midwestern United States, and defined extreme precipitation events as uninterrupted 6-h periods that surpass the 10-yr return level at a particular gauge location. They found that the majority of the rainfall that produces 6-h events accumulated in 1 or 2 h, suggesting that in some ways subdiurnal precipitation extremes play a more significant role in short-term flooding events than extremes based on daily precipitation. Building on the results of Hitchens et al. (2010), we now seek to objectively characterize the precipitating systems that generate extreme amounts of precipitation on hourly time scales.

Although flash-flood events frequently are attributed to mesoscale convective systems (Gallus et al. 2008; Schumacher and Johnson 2005, 2006), individual convective storms such as supercell thunderstorms have been shown to generate extreme rainfall (Smith et al. 2001). Such storm-type attribution is based typically on fits to conceptual models. The fits may be subjective, or employ objective criteria. Consider for example the classification scheme of Parker and Johnson (2000), who employed reflectivity criteria to delineate convective and stratiform regions within linear mesoscale convective systems (MCSs), which has been used to qualitatively characterize extreme precipitation-producing MCSs by Schumacher and Johnson (2005, 2006). They found that the majority of extreme precipitation-producing MCSs were classified as either “training line, adjoining stratiform,” “backbuilding or quasi-stationary,” or “ trailing stratiform.”

The preceding studies had objective criteria that were applied manually, thus placing practical limitations on the size of their statistical samples. This limitation can be removed through the use of automated algorithms. For example, Nesbitt and Zipser (2003) were able to quantify nearly six million “precipitation features” found in...
the Tropical Rainfall Measuring Mission satellite data collected in tropical latitudes. Similar algorithms have recently been developed for the purpose of numerical weather prediction model (NWP) verification (Davis et al. 2006a; Ebert and McBride 2000; Ebert and Gallus 2009), and also have the ability to match objects through time (Davis et al. 2006b, 2009; Skok et al. 2009, 2010). We choose the object-oriented approach developed by Baldwin et al. (2005), which will be referred to as the Baldwin Object-Oriented Identification Algorithm (BOOIA). The BOOIA forms contiguous areas of some 2D scalar into “objects.” The attributes of objects defined from rainfall data have been used to automate the classification of rainfall systems (Baldwin et al. 2005; Hitchens et al. 2010).

With the use of the BOOIA, the goal of this study is to objectively, and quantitatively, characterize the systems that produce occurrences of hourly extreme precipitation. This work is motivated in part by the desire to assess of the predictability of these occurrences, especially in light of the new generation of convective-permitting mesoscale models now in use experimentally and operationally (Kain et al. 2006, 2008; Weisman et al. 2008). For example, we are particularly interested in the system size, given the theoretical predictions of scale-dependent predictability limits (Lorenz 1969).

To determine this and other characteristics of extreme precipitation-producing systems the occurrences of extreme precipitation must first be identified. As we discuss in section 2, the rainfall data are from the stage II (ST2) precipitation product (Fulton et al. 1998), a multi-sensor product that augments radar-derived precipitation estimates with gauge measurements. Available hourly, ST2 data offer the nearly continuous spatial resolution of radar with the “ground truth” of rain gauges, making it a desirable product for use in climatological studies. In section 3, an analysis of the occurrences of extreme precipitation identified by the BOOIA from the ST2 data is presented. Two examples are provided in section 4, and conclusions and future work are presented in the final section.

2. Methods

Hourly precipitation was based on ST2 data from the beginning of the dataset (July 1996) through the end of 2010. The BOOIA was applied to each hour of ST2 data to form precipitation objects from pixels with values exceeding a user-defined threshold (Figs. 1a,b). In this case the threshold was 6 mm h⁻¹, which represents a lower bound for convective rainfall (Johnson and Hamilton 1988). The algorithm also requires a user-defined search radius, defined here as one pixel (the minimum allowed value), which has the horizontal dimensions of 4 km × 4 km. Thus, a set of pixels each meeting the rain-rate threshold and separated by no more than a single pixel comprising a single object (Figs. 1b,c). Objects of less than five pixels in size were discarded to prevent inclusion of undersampled characteristics. A sensitivity analysis is provided in the appendix, which details the impact of the choice of precipitation threshold and minimum object size. Similarly, any object with a maximum precipitation rate greater than 104 mm h⁻¹ was considered erroneous and removed from the dataset. This value was chosen because it equates to the “hail cap” on reflectivity used when deriving radar-estimated precipitation (Fulton et al. 1998). Additional quality control measures were performed to remove erroneous objects resulting from systematic errors that were identified in the dataset (R.S. Schumacher 2011, personal communication). Objects were sorted by the location of maximum precipitation, and for instances in which 100 or more objects shared the same location of maximum each object was removed. Furthermore, for instances in which at least 20 objects shared the same location of maximum precipitation, when arranged chronologically if
3 or more of these objects were identified within 24 h of the previous object they were removed.

Our focus is on precipitation occurrences in the midwestern United States, wherein the Weather Surveillance Radar-1988 Dopplers (WSR-88Ds) are well sited (and suffer minimally from issues like beam blocking due to steep orography) and hence the ST2 dataset is relatively complete. The importance of agriculture in the Midwest also supports our choice of this study area since precipitation, especially extremes, affects yields and thus has a large economic impact. We identified 365,900 precipitation objects whose pixel of maximum precipitation was located within the area bounded by 36°–47°N latitude and 80°–97°W longitude (Fig. 2). After removing objects with maxima exceeding 104 mm h\(^{-1}\) (271 objects) and performing other quality control (18,372 objects), those with maximum precipitation rate values in the 99th percentile (55.4 mm h\(^{-1}\)) were considered extreme, resulting in 3,484 occurrences of extreme precipitation. While a more statistically robust technique for identifying occurrences of extreme precipitation would be preferred, such as return levels based on the North American Regional Reanalysis (NARR) dataset (Mesinger et al. 2006), a reliable approach to identify spatial extremes has yet to be developed (Coles 2001). Other studies (Cayan et al. 1999; Gershunov and Cayan 2003) have used a quantile approach to identify precipitation extremes, and this is an improvement over choosing an arbitrary threshold. The threshold of 55.4 mm h\(^{-1}\), compares favorably with the 10-yr return levels for hourly events from Hitchens et al. (2010).

To supplement the attributes output by the BOOIA and better understand the conditions giving rise to extreme precipitation-producing systems, environmental conditions associated with each object were obtained from the NARR dataset (Mesinger et al. 2006). Environmental variables of particular interest are convective available potential energy (CAPE), precipitable water (PW), and the speed and direction of the arithmetic mean (hereafter mean) cloud-layer wind (\(V_{\text{CL}}\)). Following Corfidi (2003), the \(V_{\text{CL}}\) is defined as

\[
V_{\text{CL}} = \frac{V_{850} + V_{700} + V_{500} + V_{300}}{4}. \tag{1}
\]

The NARR data are available every 3 h, on a grid with horizontal spacing of 32 km. The environment prior to instances of extreme precipitation was evaluated by averaging values in a 3 × 3 gridpoint window centered on the NARR grid point nearest to location of maximum precipitation at the NARR analysis time valid no less than 3 h (but no more than 6 h) prior to time of the object. For example, the NARR analysis at 1800 UTC
would have been used to describe the environment of an object identified at 2300 UTC.

3. Analysis

Characteristics of the 3484 extreme precipitation objects are summarized in Fig. 3. The total horizontal area of the precipitation objects $A_o$ varied from hundreds to over 100 000 km$^2$, and the maximum 1-h precipitation rate $P_{max}$ ranged from the minimum threshold for extreme occurrences (55 mm h$^{-1}$) to the upper limit based on the WSR-88D hail cap (104 mm h$^{-1}$). The mean extreme precipitation object had an $A_o$ of slightly more than 7000 km$^2$ and $P_{max}$ of 69 mm h$^{-1}$. This mean $A_o$ value compares well to the findings of Hitchens et al. (2010); however, the mean object $P_{max}$ is much larger in this study since an object-oriented approach to defining extreme occurrences is much more likely to include the highest precipitation accumulations.

For the purpose of comparison the characteristics of both extreme and nonextreme convective precipitation occurrences are examined (Fig. 4). The largest difference between distributions is found in comparing the areal sizes of objects, with the predominance of larger “extreme” objects affirming that large systems (e.g., MCSs) are often responsible for producing heavy precipitation. Environmental variables from NARR data are also used to compare extreme and nonextreme convective precipitation. The distribution of mean cloud-layer wind speed associated with extreme occurrences was comparable to nonextreme occurrences, although the maximum was slightly slower for extreme precipitation. The distribution of extreme occurrences was similar to nonextreme occurrences for both CAPE and PW.

A receiver operating characteristic (ROC) diagram (Mason 1982) is useful in assessing the discrimination skill of a variable by comparing the probability of detection (POD) to the probability of false detection (POFD).
at a series of ordered levels. The area under a curve (AUC) on an ROC diagram is a useful measure of skill, with an AUC of 0.5 representing no skill and 0.7 considered the minimum AUC value for practical discrimination (Brooks et al. 2011). Predictably, areal size was highly skillful in discriminating between extreme and nonextreme occurrences of convective precipitation with an AUC of 0.88 (Fig. 5), while CAPE (0.62) had nominal skill, and both $V_{\text{CL}}$ (0.51) and PW (0.53) demonstrated nearly no skill. The primary factor leading to the increased discrimination skill of areal size is likely higher precipitation efficiency. Large convective systems, such as MCSs, usually experience less entrainment of dry air, reducing evaporation, and enhancing efficiency (Blyth 1993; Lamb 2001). Another factor contributing to the increased discrimination skill of object size is the affect it has on the duration of more intense rainfall (Doswell et al. 1996), where longer durations are associated with larger convective systems. Furthermore, the lack of discrimination skill of PW can be attributed to an increase in precipitation efficiency with lower (~30 mm) values (McCaul et al. 2005).

In light of the convective nature of precipitation, it is not too surprising to find that the majority of these extreme precipitation occurrences (2462) took place during the summer season (June, July, and August) similar to the findings of Maddox et al. (1979) and Brooks and Stensrud (2000), with a majority of occurrences associated with smaller objects also occurring during this season. Figure 6 shows a buildup to this summer maximum during the spring months—73 occurrences in March, 144 in April, and 374 in May—and then a significant decrease in the number of extreme occurrences between August (700) and September (203). This is likely due in part to a shift in the synoptic flow regime between the summer and autumn seasons, resulting in a decrease in moisture transport from the Gulf of Mexico, a decrease in in situ moisture from evaportranspiration, and stronger synoptic-scale flow causing faster-moving systems.

The diurnal cycle of these occurrences of extreme precipitation, shown in Fig. 7, reveals a peak during the late afternoon into the night (1900–0200 UTC), with a decline in the frequency of occurrences into the early afternoon, reaching a minimum at 1400–1500 UTC. This diurnal cycle of occurrences of extreme precipitation is likely the result of convection initiating in the mid-to late afternoon in the Great Plains region of the United States and the western portion of the area of study, then moving eastward through the study area, often in the form of MCSs (Carbone et al. 2002). Storms that initiate on the high terrain do not reach the Midwest until later in the night or early morning, likely contributing to the frequency of occurrences at these times. Further examination of the diurnal cycle across the east–west extent of the midwestern region (Fig. 8) shows that the majority of extreme occurrences take place in the western
portion of the region. In both the central and eastern regions extreme precipitation is most prevalent in the 1800–0000 UTC timeframe, while these occurrences are more frequent during 0000–1200 UTC in the western region. These results compare well to the diurnal cycle of heavy precipitation noted in previous studies (Wallace 1975; Winkler et al. 1988; Knievel et al. 2004), with areas in the western Midwest region receiving peak summer season heavy precipitation at night, while the remainder of the Midwest received it during the late afternoon. Winkler et al. (1988) suggested that locally generated nocturnal systems play an important role in this nighttime peak in heavy precipitation in the western Midwest region.

A comparison of the areal size of extreme precipitation objects by season in Fig. 9 shows that the range of object sizes during the winter are smallest, increasing from spring into summer. The median $A_o$ of objects is similar between the seasons; however, smaller precipitation objects are more likely during the summer and autumn. In terms of convective mode, this suggests that single- or multicellular storms or small MCSs are more common generators of extreme precipitation during summer and autumn when the distribution of object sizes is more likely to include smaller objects. There is also a diurnal pattern to the areal size of extreme precipitation objects, as illustrated in Fig. 10. During the afternoon (1800–0000 UTC) the median value of $A_o$ tends to be smaller, increasing from the night (0000–0600 UTC) into the early morning (0600–1200 UTC), and once again decreasing in the late morning (1200–1800 UTC). This pattern mirrors the typical evolution of storm systems in this region of the United States, with single- or multicellular storms responsible for producing more occurrences of extreme precipitation in the early afternoon and MCSs the more prevalent producer of extreme precipitation overnight. This is similar to the findings of McAnelly and Cotton (1989) in their study of mesoscale convective complexes, which are a subset of MCSs.

4. Examples

a. Example 1: 0200–0300 UTC 18 May 1997

Convection initiated in central Kansas at 2100 UTC 17 May 1997, propagating eastward and developing into isolated cells and small multicell clusters by 0000 UTC 18 May. By 0200 UTC the convective storms had spread...
to Missouri, and over the next hour a single, relatively isolated convective cell produced accumulated rainfall of 63 mm in central Missouri. In Benton County flash flooding was reported during this hour, resulting in $10,000 in property damage as well as damage to some roads and bridges (NCDC 1997). The storms progressed eastward, ultimately dissipating in the late morning. During the hour of the occurrence of extreme precipitation the BOOIA identified the precipitation object (number 9) from the ST2 data (Fig. 11), with an $A_o$ of 688 km$^2$. Five hours prior to this occurrence of extreme precipitation, there was 3660 J kg$^{-1}$ of CAPE and a PW value of 30 mm, with mean cloud-layer winds of 12 m s$^{-1}$ from 295° at the location that received extreme precipitation. This example illustrates a case in which a relatively small, isolated system produced extreme precipitation.

**b. Example 2: 1300–1400 UTC 11 June 2008**

On 10 June 2008 scattered convection initiated in eastern Kansas at 2300 UTC, forming a broken line of storms while propagating eastward. The southernmost cell in the line produced 98 mm of rainfall during the hour ending at 1400 UTC before ultimately dissipating by 1800 UTC 11 June in eastern Iowa. A flash flood was reported in Taylor County, Iowa, that caused $25,000 in property damage (NCDC 2008). The BOOIA identified object number 12 from the ST2 product during the hour of extreme precipitation (Fig. 12), with an $A_o$ of 6688 km$^2$. Five hours before the occurrence of extreme precipitation there was only 210 J kg$^{-1}$ of CAPE and 36 mm of PW, with mean cloud-layer winds of 18 m s$^{-1}$ from 239°.

**5. Conclusions**

The goal of this study was to identify and quantitatively characterize occurrences of hourly extreme precipitation in the midwestern United States, and by doing so also describe the nature of the convective systems that produce this extreme precipitation. Toward this end, the BOOIA was applied to ST2 convective systems, and those precipitation objects with maxima in the 99th percentile were considered extreme. Analysis of these occurrences revealed that the majority occurred during the summer months, with a pronounced drop in frequency between the summer and autumn.

The sizes of the extreme precipitation objects ranged from 80 km$^2$ (corresponding to single, relatively isolated cells) to 111 136 km$^2$ (corresponding to large convective systems), reaffirming the fact that although MCSs frequently produce occurrences of extreme precipitation, smaller-sized systems are not uncommon. Further investigation revealed that their characteristics differ based on the season and on the time of day that they occur. For instance, the distribution of object sizes was more likely to include smaller sizes during the summer compared to the other seasons, suggesting that many of the extreme occurrences produced by single-cell and small multicell systems occur during this season.
A comparison of the characteristics of extreme and nonextreme occurrences of convective precipitation indicated that differences exist between the two distributions. The areal size of objects exhibited the most skill of the variables examined in discriminating between extreme and nonextreme occurrences. Ostensibly this is due at least in part to enhanced precipitation efficiencies of large convective systems, as well as to the fact that a large system implies a long duration of rainfall (Doswell et al. 1996). This study provides an important step toward the goal of assessing the predictability of subdiurnal extreme precipitation events by quantifying the characteristics of both extreme precipitation and the systems that produce it. Our future research toward this goal will utilize high-resolution model simulations, using the BOOIA to compare model
output to the occurrences of extreme identified through the current study.

Acknowledgments. The authors thank Alexander Gluhovsky for his assistance with extreme value theory, and Benjamin Schwedler for providing coding assistance to plot BOOIA objects. The constructive comments and suggestions made by Russ Schumacher and the two anonymous reviewers helped improve the manuscript. Computing resources were provided by the Purdue University Rosen Center for Advanced Computing. Stage II data was provided by NCAR/EOL under sponsorship of the National Science Foundation. The National Science Foundation is acknowledged for its support of this research through Grant NSF ATM-0756624.

APPENDIX

Sensitivity Analysis

To examine the influence of user-defined BOOIA parameters such as the precipitation threshold and minimum size on characteristics of the objects identified, 100 h of ST2 data were randomly selected. The BOOIA was run using precipitation thresholds varying from 1 to 26 mm (~1 in.) and minimum size restrictions of 1, 5, and 10 pixels. The various BOOIA configurations were compared on the basis of the number of objects identified, the arithmetic mean size of all objects identified, and the 99th percentile value of the maximum precipitation of all objects identified—the same criterion used in this study to discriminate between “extreme” and “nonextreme” objects.

At the lowest allowed value of minimum object size (1 pixel) the BOOIA identifies at least twice as many objects as the other minimum object sizes at all precipitation thresholds analyzed here (Fig. A1), with the assumption that at some high threshold the three curves converge before the BOOIA fails to identify any objects. At the lowest values the choice of precipitation threshold plays a significant role in the number of objects identified by the BOOIA. Likewise, at low values the choice of minimum object size has much more impact on the total number of objects than at higher values.

As expected, the mean object sizes decrease with increasing precipitation thresholds (Fig. A2). The largest decrease occurs at lower thresholds because of an abundance of smaller hourly precipitation accumulations. It is also seen that smaller values of minimum object size result in smaller mean object sizes, with this effect diminishing at larger precipitation thresholds.

The 99th percentile value of maximum object precipitation steadily increases with increasing precipitation thresholds for each of the three minimum object size values, ultimately converging at the 24-mm threshold (Fig. A3). The three curves appear to be the same, albeit displaced by smaller precipitation threshold values as the minimum size increases. This behavior is not surprising since the population of maximum precipitation values is more dependent on object size than precipitation threshold values, such that with the addition of more small objects, the 99th percentile value will decrease.

The results of these sensitivity tests indicate that a configuration of the BOOIA with a low (high) precipitation threshold and small (large) minimum size would result in a large (small) number of objects identified, and thus a lower (higher) value for the 99th percentile of maximum object precipitation. The values used in this study were chosen for specific reasons—the 6-mm precipitation threshold is physically related to the minimum accumulation of convective precipitation over 1 h (Johnson and Hamilton 1988), and the five pixel minimum size assures that object characteristics are well...
sampled. While it is clear that the choice of BOOIA parameters influences the distributions of object characteristics, it is important to choose specific values based on other considerations.

REFERENCES


——, 2008: *Storm Data.* Vol. 50, No. 6, 1079 pp.


FIG. A3. As in Fig. A1, but for the 99th percentile value of maximum precipitation values.


