Schlie Emily (Orcid ID: 0000-0002-6029-4800)

A radar-based study of severe hail outbreaks over the contiguous United States for 2000-2011

Emily Elizabeth-Janssen Schlie
University of Illinois at Urbana-Champaign
Department of Atmospheric Sciences
3070 Natural History Building
1301 W. Green St. MC-104 Urbana, IL 61801
515-291-3166
Ejansse2@illinois.edu

Donald Wuebbles
University of Illinois at Urbana-Champaign
Department of Atmospheric Sciences

Scott Stevens
Cooperative Institute for Climate and Satellites-North Carolina

Robert Trapp
University of Illinois at Urbana-Champaign
Department of Atmospheric Sciences

Brian Jewett
University of Illinois at Urbana-Champaign
Department of Atmospheric Sciences

Key words: Severe hail, severe hail outbreaks, radar, hail proxy, climatology, MESH, hail

Abstract:

A radar-based hail climatology, with broad coverage and high resolution, is possible using the Next-Generation Weather Radar Reanalysis through application of the multi-radar multi-sensor (MRMS) algorithm, maximum expected size of hail (MESH). Using 12-years of MESH data we define a “severe hail outbreak day” and analyze the

characteristics and frequency of severe hail and severe hail outbreaks, including an analysis of hail swaths. Thresholds are set to signify severe hail in terms of MESH, and automated quality control measures are implemented. When comparing severe hail days in MESH to reports, we find a linear relationship between MESH and reports. Several case studies are also included to highlight the utility of MESH when studying outbreaks of severe hail, specifically regarding outbreak events that occur in low population areas. With the caveat that this is a relatively short time period, we find that severe hail days decrease while severe hail outbreak days increase over the period 2000-2011. The increase in outbreaks is happening primarily in the month of June, where the number of severe hail days stays fairly constant over the 12-years. This suggests that the increase in outbreaks is mainly taking place on days when severe hail already occurs. When examining hail swath characteristics we find that there are a greater number of hail swaths (with a major-axis-length (MAL) of at least 15km) on outbreak versus non-outbreak days. Additionally, hail swaths with the largest MALs occur on outbreak days.

1. Introduction

Herein we seek to better understand single-day, widespread occurrences of large or "severe" hail (hereinafter, diameter ≥ 0.75 inches), as exemplified by the events shown

in Figures. 1a and 2a-b. By virtue of their large geographical influence alone, these types of "hail outbreak" events have the potential to be relatively more impactful to society than localized events, even though localized hail events can still lead to significant damages and subsequent losses (Changnon et al., 2009). Outbreaks of severe hail can be observed in tandem with outbreaks of tornadoes (e.g., Shafer and Doswell 2010). For example, very large numbers of tornadoes (>60) were reported during the events in Figs. 1a and 2a. In contrast, the event shown in Fig. 2b was dominated by reports of hail, but also included numerous (80) non-tornadic wind reports and 25 tornado reports. Although there have been some efforts to develop flexible methods to define and characterize mixed-hazard, "severe thunderstorm" outbreaks (e.g., Shafer and Doswell 2010), these will ultimately be limited by the underlying data source.

Indeed, previous studies of the characteristics and frequency of severe hail over the United States have utilized national report-based databases that are subject to well-documented biases (Changnon, 1977; Changnon, 1999; Changnon and Changnon, 2000; Changnon, 2008; Changnon *et al.,* 2009; Kelly *et al.,* 1985; Schaefer and Edwards, 1999; Schaefer *et al.,* 2004; Doswell *et al.,* 2005; Allen and Tippett, 2015; Blair *et al.,* 2017; Barrett and Henley, 2015). A similar statement can be made for hail climatology efforts elsewhere in the world (e.g. Cao 2008; Shuester *et al*., 2005; Tuovinen *et al*., 2009). Such databases likely underestimate hail size, with the largest hailstones often going

unreported, in part because eyewitness storm reports tend to be localized to heavily populated areas (Blair *et al.,* 2017). Hail size descriptions may also be exaggerated or reported in relation to some object, such as a pea or baseball, which creates categories of size ranges rather than direct measurements (e.g. Allen et al., 2017; Doswell *et al.,* 2005; Schaefer *et al.*, 2004; Sammler, 1993). These types of biases are being exacerbated by reporting via social media and storm chasers (Allen and Tippett, 2015). In addition, spatial differences between actual events and reports can also be biased. Doswell *et al.* (2005) noted that reports are "point based" whereas the actual event covers a larger area and for a longer time.

Satellite remote sensing offers one possible substitute to reports. Several studies provide methods to detect hail-producing storms utilizing satellite measurements (Cecil *et al*., 2012; Ferraro *et al*., 2015). However, the study by Cecil *et al*. (2012) presents other limitations, such as detection not being possible during the mid-afternoon to evening, due to the sampling times of the satellite used. Additionally, this method tends to over-detect over the tropics, as noted by Ferraro et al. (2015). Ferraro *et al*. (2015) provide an alternative to the methods of hail detection used by Cecil *et al*. (2012), allowing for diurnal sampling. However, the probability of detection for this method is around 40% compared to reports. Nonetheless, satellite-based hail detection is a useful tool for regions where radar coverage and reliable reporting are sparse.

Due to the well documented biases within storm report databases, and the spatiotemporal limitations of reports, severe hail outbreaks are especially challenging to identify and difficult to study. Shafer and Doswell (2010) developed a ranking method for all types of severe weather outbreaks from 1960-2006. However, they did not seek to define an outbreak in terms of a particular hazard type, but to rank the most significant of outbreaks. An objective, automated method of analyzing severe hail outbreaks, based on spatially consistent data, is required to remove subjectivity and provide a more consistent record when determining trends and identifying other characteristics of severe hail outbreaks. Herein we propose such a method.

We also propose a method to objectively quantify hail swaths, which have been defined in various ways, typically with hail report data (e.g., Changnon et al. 2009). Alternatives to hail reports have recently been used to gain insight about the size and path of hail swaths. Basara *et al.* (2007) utilized the hail detection algorithm, outlined in Witt *et al.* (1998), and geographical information systems (GIS) tools. Their region of interest was limited geographically to the Southern Great Plains of the United States, and temporally to 2001-2003. Additionally, their method of swath detection required input from storm reports to identify hail days and manual contouring of hail swaths. Thus, no fully automated method of analyzing hail swaths in the United States currently exists.

Because the type of storms (i.e., supercells) that produce large amounts of hail tend to be long lived, with the ability to cover a considerable area (Bunkers *et al*., 2006), we hypothesize that hail swaths on outbreak days would be longer and more frequent than those occurring on non-outbreak days.

Herein we show that multi-radar composites of Next-Generation Weather Radar (NEXRAD) data can be used to obtain information regarding hail swaths, severe hail and severe hail outbreaks over the United States. Our basic approach builds on that of Cintineo *et al.* (2012), who used the Multiyear Reanalysis of Remotely Sensed Storms dataset, and multi-radar multi-sensory (MRMS) algorithms, such as maximum expected size of hail (MESH) (Witt *et al.,* 1998), to examine the presence and severity of hail over a 42-month period. The use of MRMS data mitigates single radar issues such as the "cone-of-silence", beam broadening at far ranges, and terrain blockage (Cintineo *et al.,* 2012; Smith *et al.,* 2016). While MESH was not found by Cintineo *et al.* (2012), or a recent study by Ortega (2018), to be an effective direct predictor of the maximum size of hail, it is a useful tool to verify the presence of severe hail. It also provides better spatial and temporal coverage than reports and less human effort to gain such advantages. Nisi *et al.* (2016) also demonstrated the international applicability of radar-based hail proxies, such as Maximum Expected Severe Hail Size (MESHS) and Probability of Hail, using damage reports from insurance companies for verification. The variable MESHS

differs slightly from that used in Cintineo *et al.* (2012), as the echo top height at 50 dBZ and the freezing level height are used, rather than the entire reflectivity profile. Their successful use of MESHS, along with other studies such as Soderholm et al. (2016), shows the international importance of such radar-based hail proxies in areas where reporting networks are sparse. Studies by both Cintineo *et al.* (2012) and Ortega (2018) perform verification of MESH by utilizing data from the Severe Hazards Analysis and Verification Experiment (Ortega et al., 2009). The more recent study by Ortega (2018) utilizes a longer time period for verification, from 2006-2012. Both studies show MESH as a useful option when studying historical severe hail events.

The main objectives of this study are as follows: 1) create a 12-year severe hail dataset, using the hail proxy MESH, with fully automated quality control; 2) define and analyze severe hail days and severe hail outbreak days using the 12-year MESH dataset; 3) define a hail swath and analyze hail swath characteristics on both outbreak and non-outbreak days.

The remainder of the paper is organized as follows. Section 2 outlines data and methods used, including the implementation of automated quality control. Section 3 covers the definition of a severe hail outbreak and outlines several case studies that highlight the utility of the MESH dataset. Section 4 defines a hail swath. Section 5 provides analysis of

outbreak days in terms of MESH and subsequent hail swath characteristics. Section 6 summarizes the findings and resulting conclusions.

2. Data and Methods

Data used in this study are from NOAA's NEXRAD reanalysis for the time period 2000-2011. While Cintineo *et al.* (2012) analyzed 42-months of MESH data, focusing mainly on showing the successful utility of MESH as a verification tool, this is the first study to analyze severe hail outbreaks. We utilize a much longer period of data, 12-years of the MRMS hail proxy data. Two variables are employed from the MRMS data set: MESH and Composite Reflectivity (CREF), which provides the maximum radar reflectivity in the column. MESH is derived from a thermally weighted vertical integration of radar reflectivity, from the melting level to the storm top, utilizing environmental temperature information. More information on the derivation and quality control of MESH can be found in Witt *et al.* (1998), and Lakshmanan *et al.* (2006). Additional information regarding the development and application of the MRMS severe weather products can be found in Smith *et al.* (2016) and Ortega (2018). Both MESH and CREF use multi-radar data that have been interpolated from the native radar-based (spherical) coordinate system to a uniform 0.01 x 0.01-degree latitude/longitude grid (3501 by 7001 grid points total). While MRMS MESH and CREF are now available operationally at 2-minute

intervals, the reanalysis utilized here was produced at 5-minute time intervals. Hail reports, utilized to further highlight the utility of MESH, are taken from the National Oceanic and Atmospheric Administration/National Climatic Data Center (NOAA/NCDC) Storm Events database. All daily data cover a period from 12 UTC to 12 UTC.

At each 5-min interval and for each grid point within our contiguous United States domain, we consider a point to be experiencing severe hail when MESH is at least 29 mm but less than 100 mm. Our minimum threshold is based on Cintineo *et al.* (2012), who found that MESH values ≥ 21 mm correspond best to "any hail", and that MESH values ≥ 29 mm correspond best to "severe hail" based on reports. To remove unrealistically high hail estimates, we set a maximum threshold of MESH < 100 mm. While it is not unrealistic to expect hail greater than 100 mm, setting a conservative threshold aided in eliminating erroneous hail signatures while still allowing for adequate detection of severe hail. The sum of all occurrences of severe hail over our domain (i.e. grid points), even if one grid point is "activated" more than once, for a given time period, is referred to as MESH counts. Additionally, the number of unique points with at least one MESH-based indication of severe hail for each day is referred to as "MESH area." This provides an estimate of the areal extent of severe hail over the time period.

While there are basic quality control measures in place during the generation of the MESH product (Lakshmanan *et al.,* 2006, 2007a, 2007b, 2010), some erroneous hail

signatures still appear in the data (Figure 3a). Such signatures are typically short-lived and appear for less than a day, and are possibly a result of a malfunctioning radar. Some of these MESH error signatures correspond to unrealistically high composite reflectivity values, not surprising as both are radar-reflectivity-based products. Thus, by setting a maximum CREF threshold of 80 dBZ, which exceeds the approximate reflectivity associated with "softball"-sized hail in a hail storm (Rinehart 2010), we are able to remove a large number of these errors while still retaining severe hail signatures. Specifically, if CREF exceeds this threshold, a grid point is removed for the day. Through examination of several error signatures, it was determined that additionally removing all points within approximately 40-kms, in any direction, of an erroneous data point, was necessary to remove the entire error signature associated with a single radar.

However, very apparent errors remained in the MESH dataset, even with the CREF constraint applied. Numerous five-minute time periods were found to indicate unrealistically high MESH area (e.g. 67000 grid points of MESH area). Further investigation of several of these specific events suggests that severe hail indications of more than 3000 unique points for one 5-minute interval are erroneous. Only two instances of 5-minute files with MESH area between 2000 and 5000 grid points happened on days with no very apparent error signature. Thus, five-minute periods with MESH area > 3000 grid points were removed from the dataset, 3000 grid points being

chosen as a conservative threshold. Figure 3b demonstrates that application of these additional quality control measures successfully removes apparent error signatures, while keeping real signatures intact.

Even with the additional quality control measures described herein, there are still several caveats associated with utilizing a radar-based proxy. While the issue of beam broadening is minimized in many areas due to the use of multiple radars, there are still some areas in the US where there is only single radar coverage. One implication is that at large radar ranges, where the resolution volume of the radar is large, the area of hail could be spatially overestimated (e.g., Cintineo *et al.,* 2012). Some of this can be effectively filtered out utilizing the quality control measures described in this section, however, it is possible for some false hail detection to remain. Cintineo *et al.* (2012) also noted that overestimation of hail size is also possible in areas of single radar coverage, when beam-filling occurs for an elevated beam height, filling the resolution volume with high reflectivity values. Finally, Bunkers and Smith (2013) suggest that an underestimation of hail size, and perhaps an overestimation of hail coverage, are possible if the melting level is below the lowest radar beam.

3. Defining a severe hail outbreak with MESH

To employ MESH as a tool for identifying severe hail outbreaks, we first compare it against reports of severe hail. In general, when considering all severe hail, not just outbreaks, there is a strong relationship between the spatial extent of the MESH product and the number of storm reports that are received on a given day. Figure 4 shows the paired relationship between MESH area and the number of severe hail reports received for each day in the 12-year dataset, on which there was at least one indication of severe hail in the radar-based dataset. While outliers exist due to reasons that we will demonstrate next, there is a clear linear relationship between these values, with an R^2 value of 0.60.

It is instructive to examine two pairs of days from the dataset to illustrate the advantage of an objective measure of hail occurrence and highlight examples that would largely be accepted as an outbreak either based on storm reports or MESH. Table 1 lists these event dates along with the associated MESH area for that day and the number of severe hail reports in the Storm Events Database.

Looking at Figure 1 we show an example of two days that, according to the MESH-based analysis, were very similar in areal extent of hail (bottom two rows of Table 1). The 12 March 2006 event (MESH area = 15537 grid points) occurred in the central US, and covered several large metropolitan areas, including Kansas City, MO. The result is a

higher concentration of people witnessing the event and reporting it (594 reports received). In contrast, the 22 July 2011 event (MESH area = 15458 grid points) appears to have occurred almost entirely over rural areas, which we argue can explain why only 93 reports were received, nearly five times fewer than that of the 12 March case (e.g., see Trapp *et al.* 2006).

A second pair of days illustrates a similar scenario (Figure 2, top two rows of Table 1). On both 10 April 2009 and 15 June 2009, a similar number of hail reports were received (387 and 386, respectively). However, on the former, we find a MESH area of only 4947 grid points, a relatively small event. The event spanned several cities including Atlanta, GA, and major highways such as Interstates 20 and 85 through Georgia, resulting in a much higher number of reports. On the latter, a MESH area of over 23000 grid points indicates a much larger, widespread event. However, as before, this event occurred mostly in rural areas, resulting in a lower number of reports than may be expected from an event of such magnitude. In fact, the 15 June case showed a radar-indicated hail extent of about 1.5 times larger that of the 12 March case from the previous pair, yet resulted in in around 200 fewer reports. This further highlights the effect population density has on hail reports.

These cases illustrate the benefits of a radar-based analysis of severe hail when studying outbreaks, namely, that it removes the population dependency of subjective reports and provides a consistent estimate of hail regardless of where it falls. Encouraged by these benefits, our next task is to define a severe-hail outbreak criterion using MESH. We assume that an outbreak is a relatively rare, widespread event, and thus quantify outbreaks using a MESH area that represents a comparably low occurrence frequency. Specifically, we use a MESH area threshold of 6000 grid points, which is the 90th percentile of the distribution of all MESH area values in the 12-year period. These points are not required to be continuous, implying that the MESH area can contain gaps. This does result in some uncertainty in terms of whether all points comprising the outbreak were caused by the same synoptic/mesoscale forcing, but such a requirement would result in the loss of significant events. Note that if the threshold is increased to 10000 grid points (the ~95th percentile value), the number of outbreaks per year decreases, but nearly identical trends are found. Thus, in section 5, we will use the 90th percentile criterion to quantify the occurrence frequency and other characteristics of severe hail outbreaks.

4. Defining a hail swath with MESH

As alluded to in section 1, one of our hypotheses is that severe-hail outbreaks comprise relatively long hail swaths. Note that a hail swath is generally considered to be a large area of relatively contiguous hail fall. Previous studies have proposed objective criteria

for identifying a hail swath (Changnon *et al.,* 1967; Schleusener, 1966). These studies were forced to rely on storm reports in the absence of a remote sensing-based method of locating hail. Here, contiguous areas of hail fall can be easily identified within the MESH data due to its high spatial and temporal resolution.

A hail swath is defined here as a contiguous area of hail fall, occurring within the same day (24-hour period). To identify hail swaths within the MESH data we employ a Python toolkit called scikit-image, which employs connected component labeling to find continuous "objects" within the daily severe MESH dataset. Objects are grid points grouped together using "8-connectivity", meaning grid points are connected by both their faces and edges. As we are concerned with storms that produce hail over a broad area, we consider only those objects with a major-axis-length (MAL) of at least 15 km; recalling that the MESH data are on a 0.01-degree grid, which equates to approximately 1-km spacing, any reference to length or distance of hail swaths in km are approximated to the nearest 1 km. Figure 5 shows examples of some of the hail swath objects for the outbreak case of 18 April 2002. In section 5, we will quantify MAL along with several other hail swath variables, and then relate these to severe hail outbreaks.

5. Objectively defined severe hail outbreaks and hail swaths during 2000-2011

We consider a day to be a "severe hail day" when MESH area ≥ 100 (i.e. 100-km^2). Applying this threshold helps to minimize any potential remaining erroneous MESH data. It is not logical to assume that severe hail would be isolated to one single 1-km^2 grid point, for an entire day, so simply requiring MESH area to be greater than zero was not an acceptable threshold in this scenario. However, there is approximately an 89% chance of at least one eye-witness severe hail report on days when MESH area ≥ 100 grid points, when considering the entire 12-years of data. Table 2 shows average severe hail days based on MESH, severe hail outbreak days based on MESH, and severe hail days based on SPC reports, by month. There is very good agreement, for every month, between average severe hail days in terms of MESH area ≥ 100 grid points and reports > 1. However, this relationship weakens as values of MESH area and/or reports increase (Figure 4). The case studies presented in Figures 1 and 2 highlight potential examples of why MESH and reports might become less correlated as values increase, and why MESH has potential advantages over utilizing reports to study outbreak events.

Figure 6 shows maps of total severe hail counts, in terms of MESH, accumulated for each year. The most notable feature of Figure 6 is the clear boundary separating the western

US from the eastern US in terms of severe hail. The majority of severe hail is located in the Central Plains region, centered roughly on Kansas and Oklahoma. Additionally, there are a few remaining error signatures, specifically in the western US, where radar coverage is sparser, and for example, unsuppressed clutter from terrain can result in radar errors. The most notable of these error features occurs regularly in Nevada.

While there are at least some severe hail days per month throughout each year, in terms of both MESH and reports, we find that outbreaks are rare to non-existent in the fall and winter months (Table 2). The spring and summer months are the most active period for outbreaks, with over one third of June days indicating outbreaks. This is consistent with the results of Doswell *et al.* (2005), who found that non-tornadic severe thunderstorm frequency peaks in the late spring.

Figure 7 shows an annual time series (2000-2011) for outbreak days and all severe hail days, for the entire contiguous United States. This is presented with the caveat that the data record is relatively short and thus that any apparent trends (and statistical tests thereof) are sensitive to outliers. This includes the positive trend in the number of outbreak days, which have nearly doubled over the 2000-2011 period; for reference, linear regression yields an R-value of 0.65, a P-value of 0.022, and the 95% confidence interval of the slope is 0.32 to 3.33. However, as previously mentioned, short time series

are more sensitive to outliers, such as the year 2011, which was a particularly active year for hail events. As a point of reference, if we remove 2011 from the time series the P-value increases to 0.102 and the R-value decreases to 0.518. An apparent decrease in the number of days with severe hail is shown, specifically in the last few years of the time series. There is a dip in severe hail days in 2004, which is due to a lack of MESH data for October and November of that year. The data for those months in 2004 were missing. This did not impact outbreaks, as outbreaks are rare to non-existent during these months, based on other years. Figures 8 and 9 show the outbreak and severe hail day time series (2000-2011) for the warm (April – September) and cold (October – March) seasons respectively. The increase in outbreak days is isolated to the warm season, with an R-value of 0.66, a P-value of 0.018, and the 95% confidence interval of the slope is 0.39 to 3.28. The less apparent decrease in severe hail days is isolated to the cold season.

While this is based on a relatively short period, these time series suggest that over the last decade, days when severe hail occurs are decreasing while there is an increase in the number of outbreak days. However, the increase in outbreak days is isolated to the warm season, where the overall number of severe hail days show no apparent trend. The increase in severe hail outbreak days during the warm season is similar to the results from Brooks *et al.* (2014), who found a positive trend in tornado outbreak days. However, total number of tornado days shows no overall change, while severe hail days

do show a slight decrease during the cold season. Finding increases in both tornado and severe hail outbreak days is not unexpected given the similarity in storm morphology between tornado and severe hail producing storms.

When we break down the same two variables from Figure 7 by month (Figure 10) we see that while there are a considerable number of outbreak days in May and July, and a spike in April at the end of the time series, the increase is mainly occurring in June. There is no concurrent increase in overall severe hail days in June; that line remains fairly constant. This is also consistent with Brooks *et al.* (2014), and suggests days that already produce severe hail are increasingly becoming outbreak days. Again, it is important to remember the dip in severe hail days during October and November of 2004 is due to the missing MESH data from that year for those months.

We also examine outbreaks in terms of the characteristics of the hail swaths that comprise them, and then compare these hail-swath characteristics with those occurring on non-outbreak days. Figure 11 shows histograms of the number of hail swaths occurring on outbreak and non-outbreak days for the entire 12-year period. Hail swaths on outbreak days have a fairly Gaussian distribution with most days having somewhere between 20 and 35 hail swaths, with as many as 87 hail swaths on a single day. Days without outbreaks most frequently only contain between 0 to 5 swaths. It is quite

apparent that hail swaths with a MAL of at least 15-km are far more frequent on outbreak days than on non-outbreak days when severe hail occurred. It should be noted that the number of hail swaths per day is somewhat dependent on the choice of the minimum MAL in the definition of a hail swath. Additionally, Figure 12 shows an apparent positive trend in the number of hail swaths per year on outbreak days. There is also a negative trend in number of total hail swaths per year on non-outbreak days, although it is less apparent. These trends mirror what is shown in the annual time series of outbreak and severe hail days (Figure 7). It is logical to expect that if the overall number of outbreak days per year is increasing that hail swaths would become more numerous on those days.

Figure 13 shows that the most frequently occurring MAL of hail swaths is around 15 km, regardless of outbreak day status. It is somewhat expected to get the smallest allowed MAL as the most frequently occurring. However, MAL can reach into the 500 km to 600 km range on outbreak days, although those occurrences are quite rare. These more extreme MALs are nonexistent on non-outbreak days. Based on Figure 13b, an increase in the minimum MAL, in the definition of a hail swath, would result in fewer hail swaths per day overall, and exaggerate the difference in magnitude of hail swaths per day between outbreak and non-outbreak days. Regardless of outbreak or non-outbreak day, total area of hail swaths is most often between 50 to 100 grid points (km^2).

The main difference in hail swath characteristics between an outbreak day and non-outbreak is the number of hail swaths and their MAL. Far more hail swaths occur on outbreak days, and the occurrence of swaths with the largest MALs is limited to outbreak days. This suggests that the largest long track storms that consistently produce large amounts of severe hail generally occur on outbreak days, which is to be expected.

6. Conclusions

This study highlights the overall utility of a radar-based hail proxy to quantify occurrences of severe hail outbreaks. A 12-year (2000-2011) severe hail dataset was developed using the maximum estimated size of hail (MESH) radar product, with automated quality control measures applied to eliminate erroneous severe hail signatures. Additionally, definitions were provided of a severe hail outbreak and hail swath in terms of MESH.

While results show that the linear relationship between MESH area and reports weakens with increasing MESH area and reports, selected cases show MESH more consistently captures severe hail outbreaks compared to reports, regardless of the population of the

region affected by an outbreak. Additionally, the relationship between MESH and reports, for smaller hail days, could potentially be utilized when interpreting severe hail reports and subsequent biases in reports.

While precautions should be taken when interpreting the results of a short timeseries with potential outliers, there is an apparent increase (R-value of 0.65, a P-value of 0.022) in the number of severe hail outbreak days occurring over the contiguous United States, between the years 2000-2011, in terms of MESH area. This increase is limited to the warm season. There is also a decrease in the number of overall severe hail days, though not during the peak severe hail months, and the decrease is less apparent. These findings are similar to the results of Brooks *et al.* (2014) where an increase in tornado outbreaks was observed but no concurrent increase in overall tornado days was found. The increase in severe hail outbreaks is mainly occurring during the month of June and there is no concurrent increase in overall severe hail days during that month. This suggests that the added outbreak days are happening on days, in the warm season, where severe hail already generally occurred. The authors speculate that this could point to an increase in environmental conditions supportive of severe hail outbreaks, perhaps through an increase in the spatial extent of environmental conditions conducive to severe hail.

There are two main differences between outbreak days and non-outbreak days in terms of hail swath characteristics. The number of hail swaths that generally occur on an outbreak day is around five times larger than the number of hail swaths commonly occurring on non-outbreak days. Additionally, outbreak days contain the longest hail swaths, based on MAL, suggesting that long-track, severe hail producing storms tend to favor outbreak days.

Further analyses will be generated using this dataset, especially as more years of data become available. The authors are currently working on linking the MESH outbreaks to their meteorological environments using the North American Regional Reanalysis in order to analyze longer term historical trends in severe hail outbreaks.

Acknowledgments

Support for this research was provided to the first author (EJS) by the NOAA Cooperative Institute for Climate and Satellites (CICS) under sub-award CICS-NC 2014-2918-04. The authors would like to thank the National Centers for Environmental Information, CICS, and the National Severe Storms Laboratory for the development of the NEXRAD Reanalysis. The authors would also like to thank Dr. Kimberly Hoogewind

for her assistance with the development of this research, and the four reviewers for their helpful comments and questions on the manuscript.

References

Allen JT, Tippett MK. 2015. The characteristics of United States hail reports: 1955–2014. *Electronic J. Severe Storms Meteor* **10 (3)**: 1–31.

Allen JT, Tippett MK, Kaheil Y, Sobel AH, Lepore C, Nong S, Muehlbauer A. 2017. An Extreme Value Model for U.S. Hail Size. *Mon. Wea. Rev.* **145:** 4501-4519. https://doi.org/10.1175/MWR-D-17-0119.1

Barrett BS, Henley BN. 2015. Intraseasonal variability of hail in the contiguous United States: relationship to the Madden-Julian Oscillation. *Mon. Wea. Rev.* **143:** 1086-1103. https://doi.org/10.1175/MWR-D-14-00257.1

Basara JB, Cheresnick DR, Mitchell D, Illston BG. 2007. An analysis of severe hail swaths in the Southern Plains of the United States. *Trans. GIS* **11:** 531–554.

Blair SF, Laflin JM, Cavanaugh DE, Sanders KJ, Currens SR, Pullin JI, Cooper DT, Deroch DR, Leighton JW, Fritchie RV, Mezzeul II, MJ, Goudeau BT, Kreller SJ, Bosco JJ, Kelly CM, Mallinson HM. 2017. High-Resolution Hail Observations: Implications for NWS Warning Operations. *Wea. Forecasting*. **32(3)**: 1101-1119. https://doi.org/10.1175/WAF-D-16-0203.1

Bunkers MJ, Hjelmeflt MR, Smith PL. 2006. An Observational Examination of Long-Lived Supercells. Part I: Characteristics, Evolution, and Demise. *Wea. Forecasting*. **21:** 673-688.

Bunkers MJ, Smith PL. 2013. Comments on "An Objective High-Resolution Hail Climatology of the Contiguous United States". *Coorespondence, Wea. And Forecasting* **28**: 915-917. DOI: 10.1175/WAF-D-13-00020.1

Brooks HE, Carbin GW, Marsh PT. 2014. Increased variability of tornado occurrence in the United States. *Science* **346**: 349-352.

Cao Z. 2008. Severe hail frequency over Ontario, Canada: Recent trend and variability. *Geophys. Res. Lett.* **35** L14803. DOI: 10.1029/2008GL034888.

Cecil DJ, Blankenship CB. 2012. Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. *J. Clim.* **25:** 687–703.

Changnon SA, Schickedanz P, Danford H. 1967. Hail patterns in Illinois and South Dakaota. Preprints, 5th Conf. on Severe Local Storms, *Amer. Meteor. Soc.* 325-335. St. Louis, MO, USA.

Changnon SA. 1977. The scales of hail. *J. Appl. Meteor.* **16:** 626–648.

Changnon SA. 1999. Data and approaches for determining hail risk in the contiguous United States. *J. Appl. Meteor.* **38:** 1730–1739.

Changnon SA, Changnon D. 2000. Long-term fluctuations in hail incidences in the United States. *J. Climate* **13:** 658–664.

Changnon SA. 2008. Temporal and spatial distributions of damaging hail in the continental United States. *Phys. Geography* **29:** 341–350.

Changnon SA, Changnon D, Hilberg SD. 2009. Hailstorms Across the Nation: An Atlas about Hail and its Damages. *Illinois State Water Survey* 95pp.

Cintineo JL, Smith TM, Lakshmanan V, Brooks HE, Ortega KL. 2012. An Objective High-Resolution Hail Climatology of the Contiguous United States. *Wea. and Forecasting* **27**: 1235-1248. DOI: 10.1175/WAF-D-11-00151.1

Doswell CA, Brooks HE, Kay MP. 2005. Climatological Estimates of Daily Local Nontornadic Severe Thunderstorm Probability for the United States. *Wea. and Forecasting* **20**: 577-595.

Ferraro R, Beauchamp J, Cecil DJ, Heymsfield G. 2015. A prototype hail detection algorithm and hail climatology developed with the Advanced Microwave

Sounding Unit (AMSU). *Atmos. Res.* **163:** 24–35, doi:10.1016/j.atmosres.2014.08.010

Kelly DL, Schaefer JT, Doswell III CA. 1985. Climatology of nontornadic severe thunderstorm events in the United States. *Mon. Wea. Rev.* **113:** 1997–2014.

Lakshmanan V, Smith TM, Hondl K, Stumpf GJ, Witt A. 2006. A real-time, three-dimensional, rapidly updating, heterogeneous radar merger technique for reflectivity, velocity, and derived products. *Wea. Forecasting*. **21:** 802–823.

Lakshmanan V, Fritz A, Smith TM, Hondl K, Stumpf G. 2007a. An automated technique to quality control radar reflectivity data. *J. Appl. Meteor. Climatol.* **46:** 288–305.

Lakshmanan V, Smith TM, Stumpf GJ, Hondl k. 2007b. The Warning Decision Support System–Integrated Information. *Wea. Forecasting*. **22:** 596–612.

Lakshmanan V, Zhang J, Howard K. 2010. A technique to censor biological echoes in radar reflectivity data. *J. Appl. Meteor. Climatol.* **49:** 435–462.

Nisi L, Martius O, Hering A, Kunz M, Germann U. 2016. Spatial and temporal distribution of hailstorms in the Alpine region: a long-term, high resolution, radar-based analysis. *Q. J. R. Meteorol. Soc.* **142:** 1590–1604.

Ortega KL, Manross KL, Scharfenberg KA, Witt A, Kolodziej AG, Gourley JJ. 2009. The Severe Hazards Analysis and Verification Experiment. *Bull. Amer. Meteor. Soc*. **90:** 1519–1530.

Ortega KL. 2018. Evaluating Multi-Radar, Multi-Sensor Products for Surface Hailfall Diagnosis. *E-Journal of Severe Storms Meteorology*. **13:** 1-36.

Rinehart R. 2010. Radar for Meteorologists, Rinehart Publications, Nevada, Missouri, 482 pp.

Sammler WR. 1993. An updated climatology of large hail based on 1970–1990 data. Preprints, 17th Conf. on Severe Local Storms. *Amer. Meteor. Soc.* 32–35. St. Louis, MO, USA.

Schaefer JT, Edwards R. 1999. The SPC tornado/severe thunderstorm database. Preprints, 11th Conf. on Applied Climatology, *Amer. Meteor. Soc.*, 603–606. Dallas, TX, USA.

Schaefer JT, Levit JJ, Weiss SJ, McCarthy DW. 2004. The frequency of large hail over the contiguous United States. Preprints, 14th Conf. on Applied Climatology, *Amer. Meteor. Soc.*, 3.3. Seattle, WA, USA.

Schleusener RA, 1966. Project Hailswath: Final Rept., Vol.1: Summaries and recommendations. NSF Contract C-461, Rept. 66-9, *Inst. Atmos. Sci.*, South Dakota School of Mines and Technology, Rapid City, USA, 33pp.

Schuster SS, Blong RJ, Speer MS. 2005. A hail climatology of the greater Sydney area and New South Wales, Australia. *Int. J. Climatol.* **25**: 1633–1650. DOI:10.1002/joc.1199

Shafer CM, Doswell III CA. 2010. A multivariate index for ranking and classifying severe weather outbreaks. *Electronic J. Severe Storms Meteor.* **5 (1)**: 1–39.

Smith TM, Lakshmanan V, Stumpf GS, Ortega KL, Hondl K, Cooper K, Calhoun KM, Kingfield DM, Manross KL, Toomey R, Brogden J. 2016. Multi-radar multi-sensor (MRMS) severe weather and aviation products initial operating capabilities. *Bull. Amer. Meteor. Soc*. 1617-1630. DOI:10.1175/BAMS-D-14-00173.1

Soderholm JS, McGowan H, Richter H, Walsh K, Weckwerth TM, Coleman M. 2017. An 18-year climatology of hailstorm trends and related drivers across southeast Queensland, Australia. *Q. J. R. Meteorol. Soc.* **143(703):** 1123-1135.

Trapp RJ, Wheatley DM, Atkins NT, Przybylinski RW, Wolf R. 2006. Buyer beware: Some words of caution on the use of severe wind reports in a postevent assessment and research. *Wea. Forecasting* **21:** 408-414.

Tuovinen J, Punkka A, Rauhala J, Hohti H. 2009. Climatology of Severe Hail in Finland: 1930-2006. *Mon. Wea. Rev.* **137**: 2238-2249. DOI: 10.1175/2008MWR2707.1

Witt A, Eilts MD, Stumpf GS, Johnson JT, Mitchell ED, Thomas KW. 1998. An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting* **13:** 286–303.

Captions

Table 1: Daily MESH area and official severe hail reports for 4 case studies. Top two show cases with similar number of reports and very different MESH area. Bottom two cases show similar MESH areas and very different total number of reports.

Table 2: Average annual number of days with severe hail based on MESH, severe hail based on reports, and severe hail outbreaks based on MESH (2000-2011). October and November of 2004 were omitted from the average due to corrupt data during that year. The R value for days with severe hail based on MESH versus reports is 0.974.

Figure 1. a) Severe hail reports for 12 March 2006, b) Severe hail reports for 22 July 22 2011, c) MESH occurrences for 12 March 2006, d) MESH occurrences for 22 July 22 2011.

Figure 2. As in Fig. 1, except for (a)-(c) 10 April 2009, and (b)-(d) 15 June 2009.

Figure 3. a) A composite of all MESH counts for 2005 with no additional quality control. Red boxes highlight the regions where erroneous MESH signatures appear. b) A composite of all MESH counts for 2005 with the additional quality control measures applied.

Figure 4. Daily number of severe hail reports (2000-2011) and the corresponding MESH area for days when MESH area > 0. The blue line is a fit from a linear regression with an R^2 value of 0.60.

Figure 5. Hail swaths, with a major-axis-length ≥ 15 km, for outbreak day, 18 April 2002. Each color shade is a separate hail swath. Note that more swaths occurred on this day than are indicated in the map area.

Figure 6. Total accumulated MESH counts for each year in the 12-year time period. Each map is a separate year.

Figure 7. Severe hail day (top-green) and outbreak day (bottom-blue dashed) time series for the period 2000-2011. The red lines are linear trend lines fit to the data. P and R-values given for each time series in red text.

Figure 8. Severe hail day (top-green) and outbreak day (bottom-dashed blue) time series for the warm season (April – September) for the period 2000-2011. The red lines are linear trend lines fit to the data. P and R-values given for each time series in red text.

Figure 9. Severe hail day (top-green) and outbreak day (bottom-dashed blue) time series for the cold season (October-March) for the period 2000-2011. The red lines are linear trend lines fit to the data. P and R-values given for each time series in red text.

Figure 10. The blue lines (bottom - dashed lines) are outbreak days and the green lines (top lines) are severe hail days. Shown by month for all 12 months of the year between the years 2000-2011.

Figure 11. a) The number of hail swaths between 2000-2011, during outbreak days. Bin values for 15-20 and 20-25 are 52 and 99, respectively. b) The number of hail swaths occurring between 2000-2011 on non-outbreak days when severe hail occurred.

Figure 12. Blue line is the total number of hail swaths per year on outbreak days for the time period. The red line is a linear trend line fit to the data. R and P values are 0.683 and 0.014 respectively.

Figure 13. a) A cumulative distribution function of the Major-axis-length of hail swaths from 2000-2011 for hail swaths occurring on both outbreak (blue dashed line) and non-outbreak (solid green line) days. b) Line graph showing the frequency of the MAL of hail swaths from 2000-2011, for outbreak days (blue dashed line) and non-outbreak (green solid line) days.

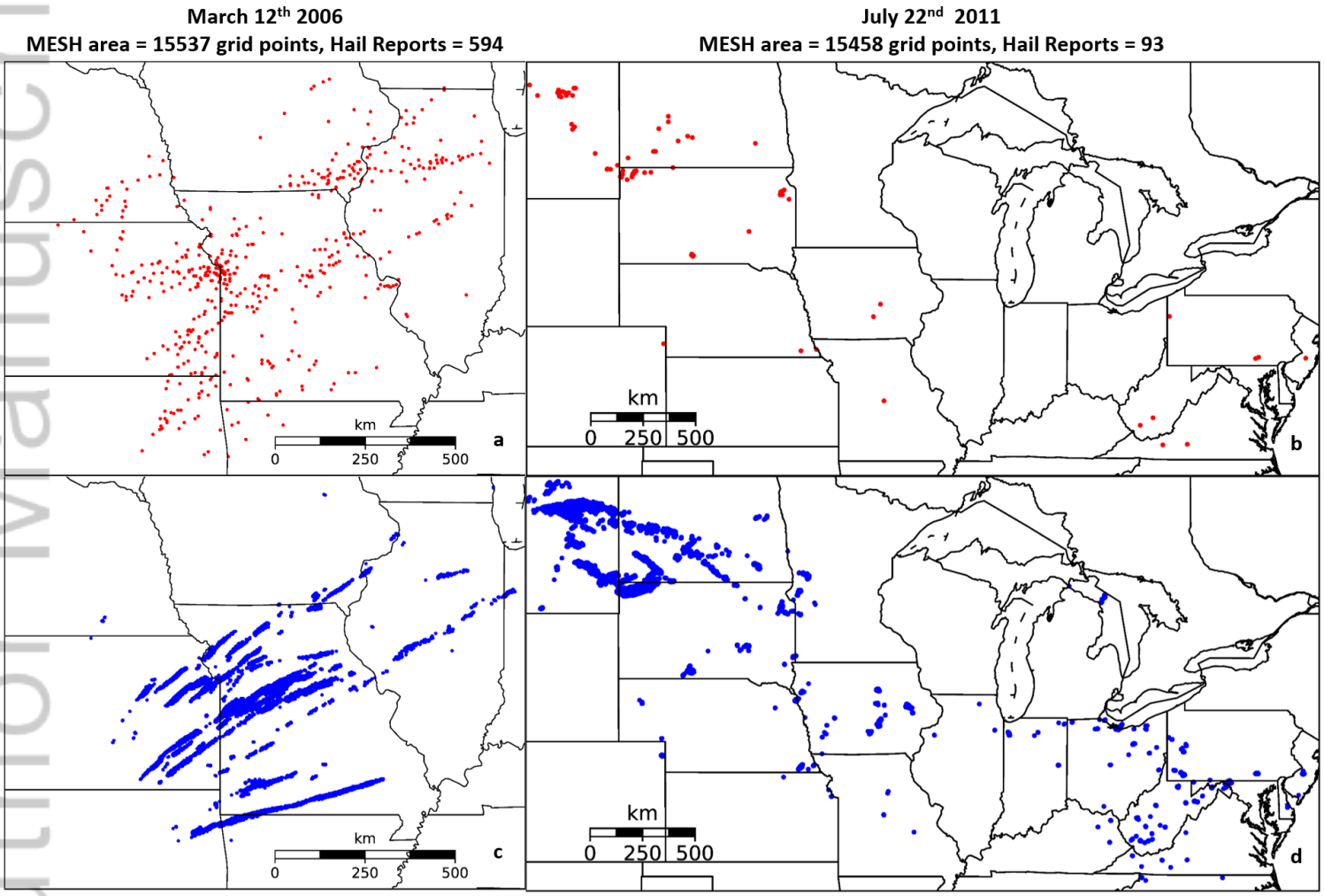


Figure1.tif

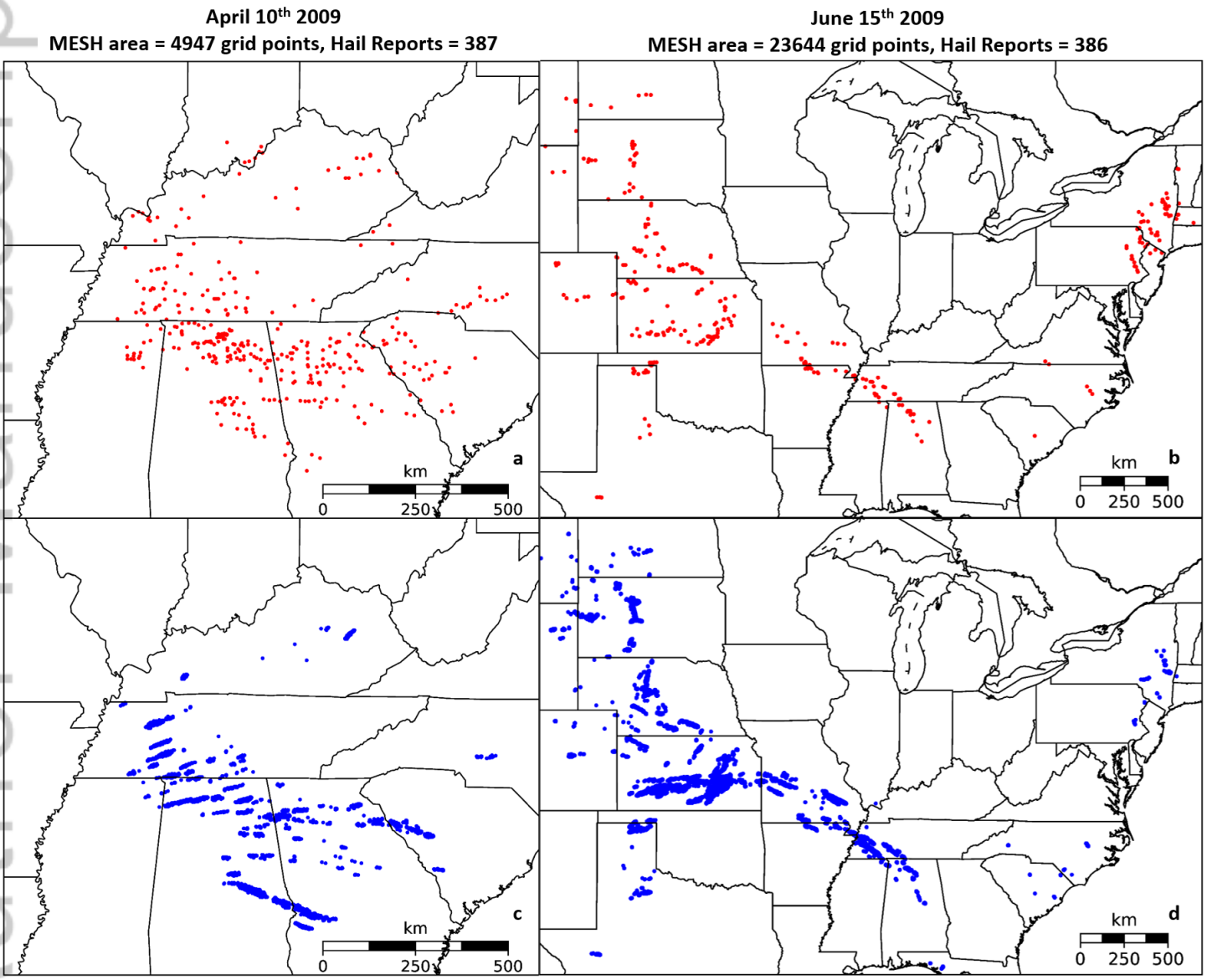


Figure2.tif

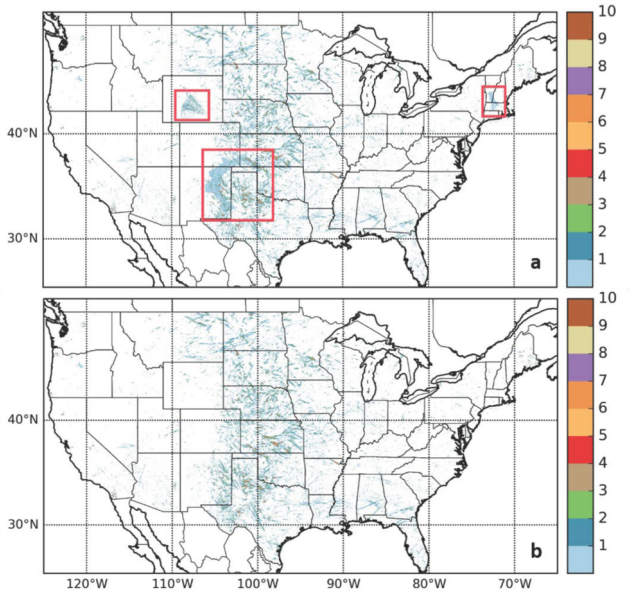


Figure3.tif

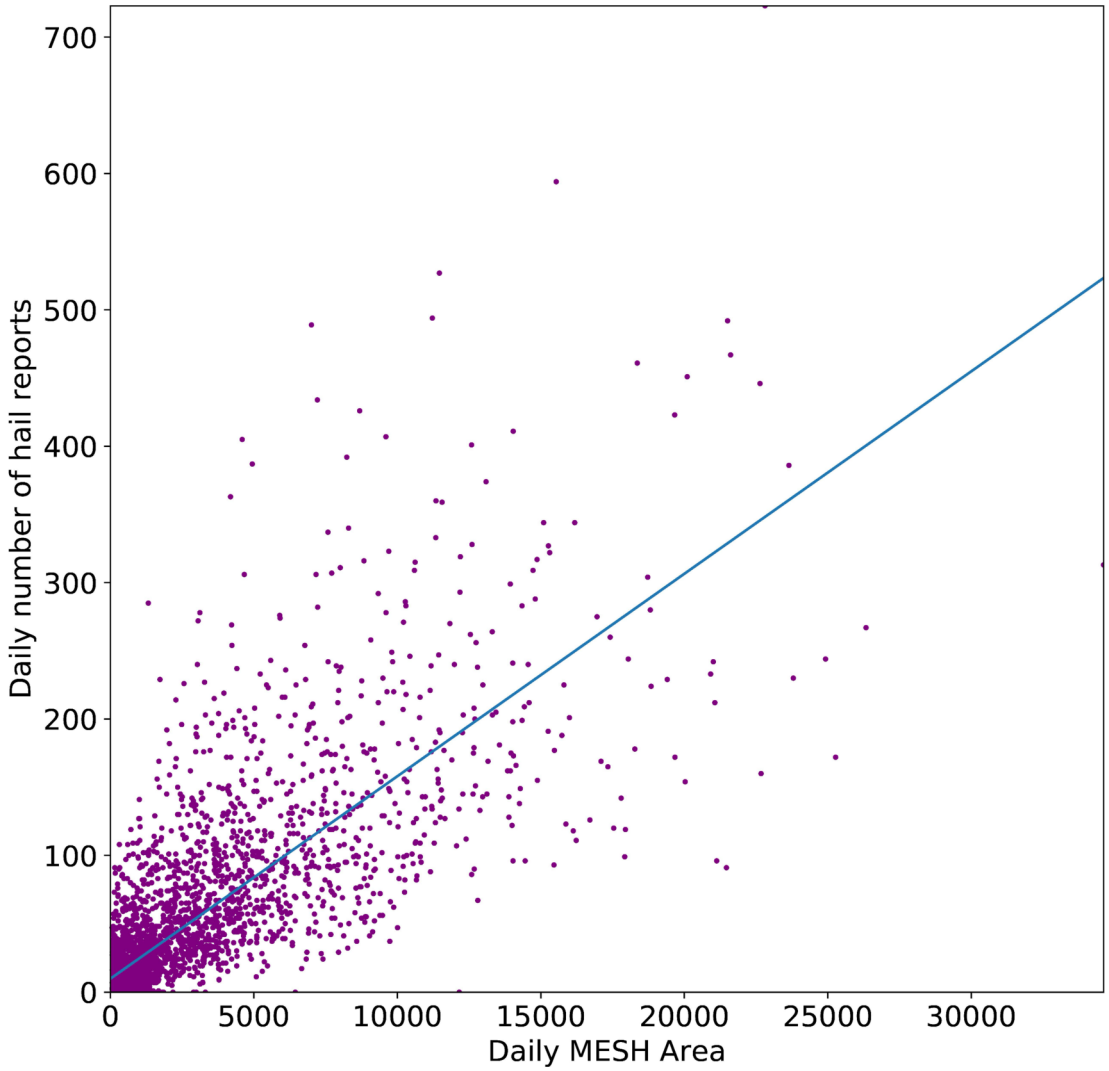


Figure4.tif

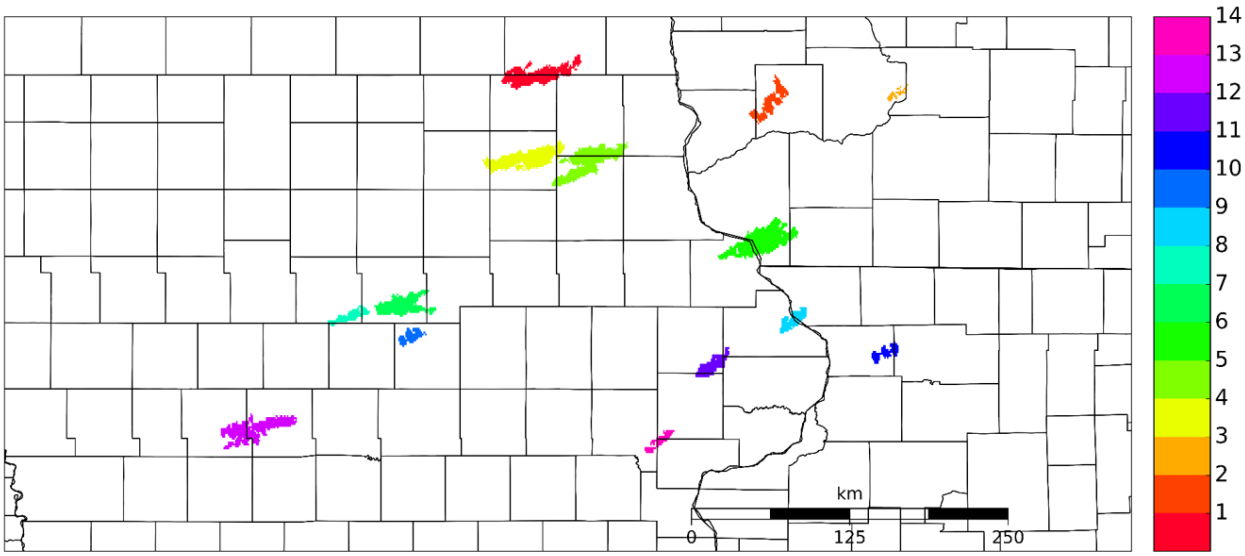


Figure5.tif

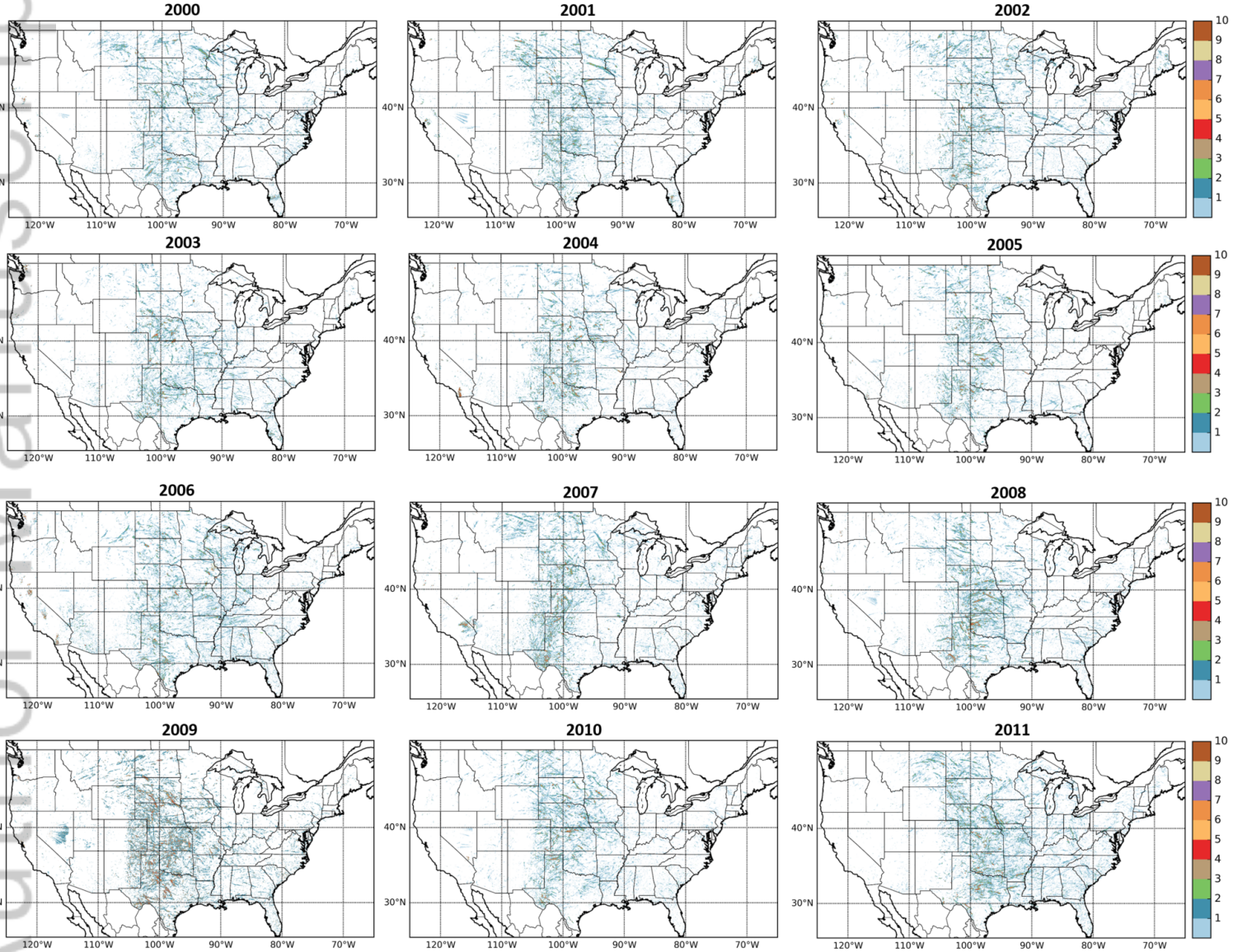


Figure6.tif

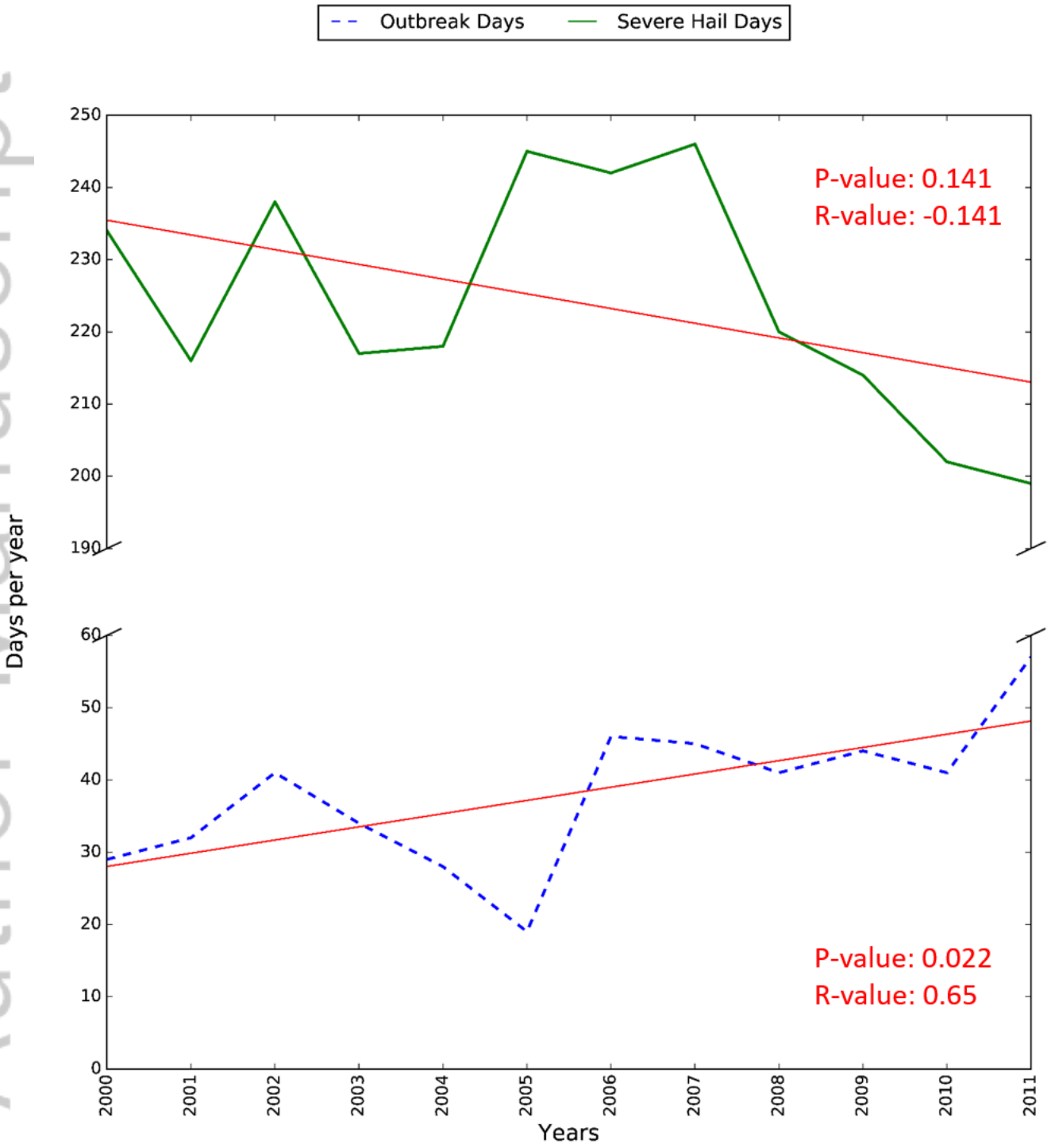


Figure7.tif

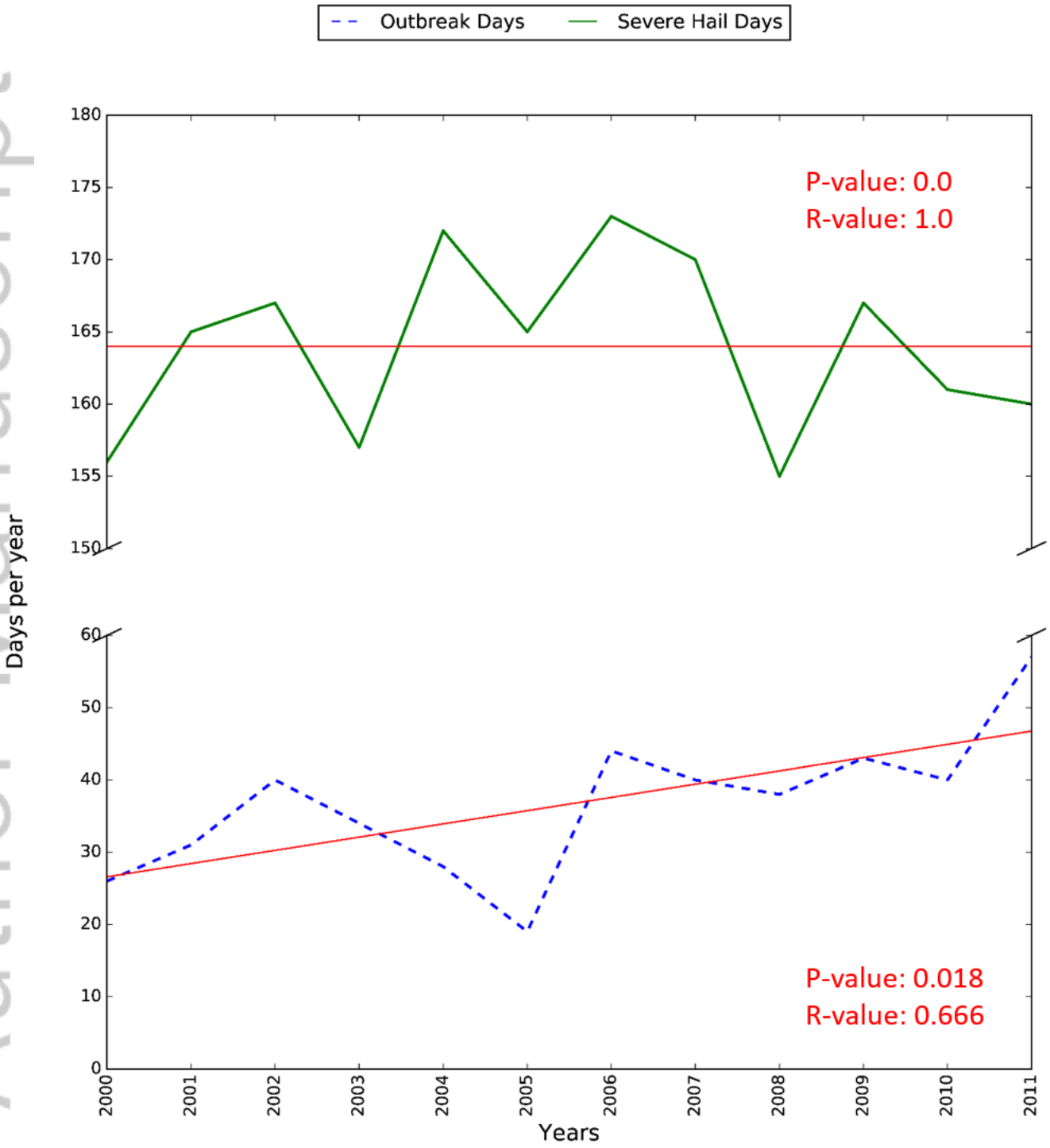


Figure8.tif

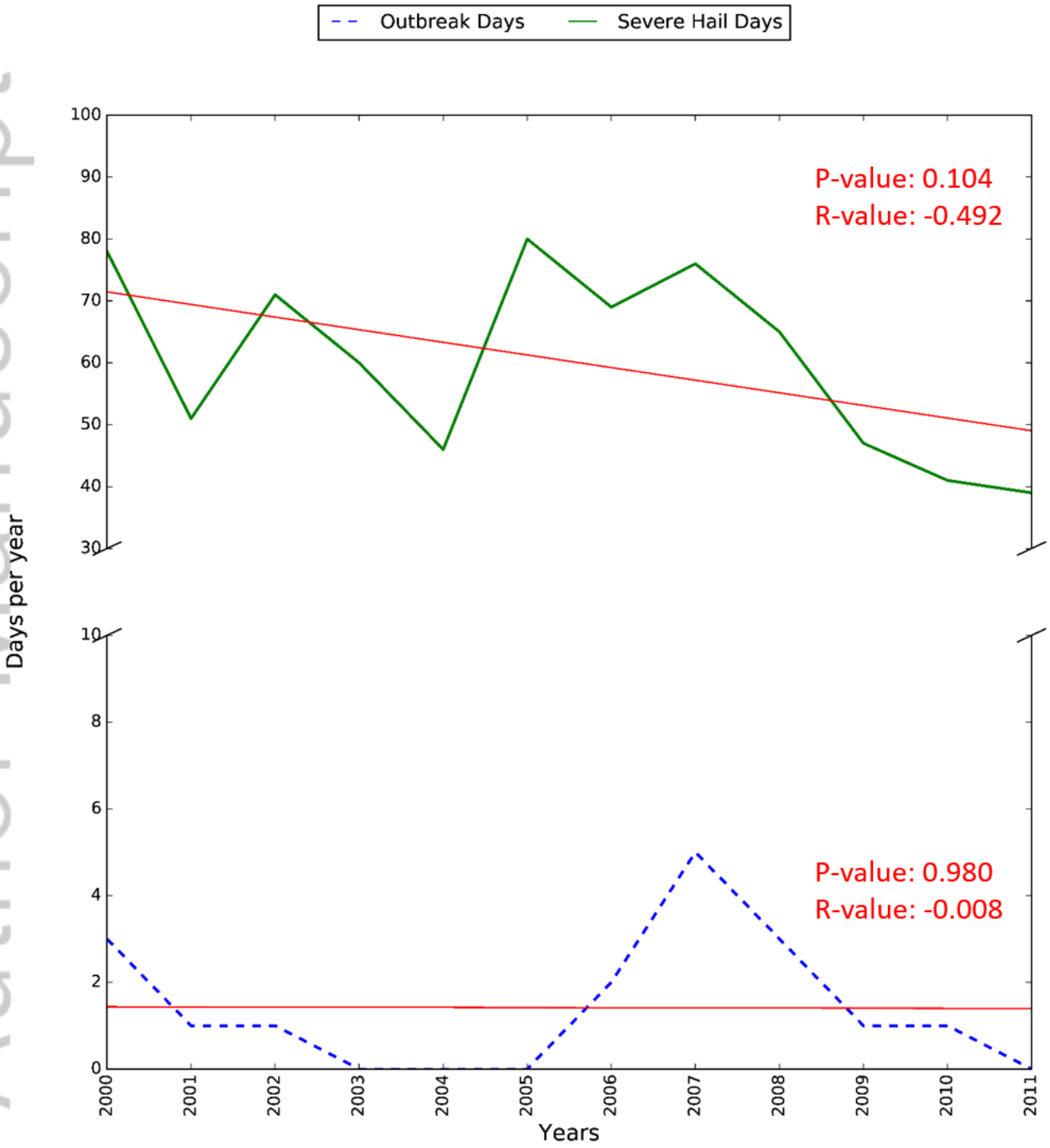


Figure9.tif

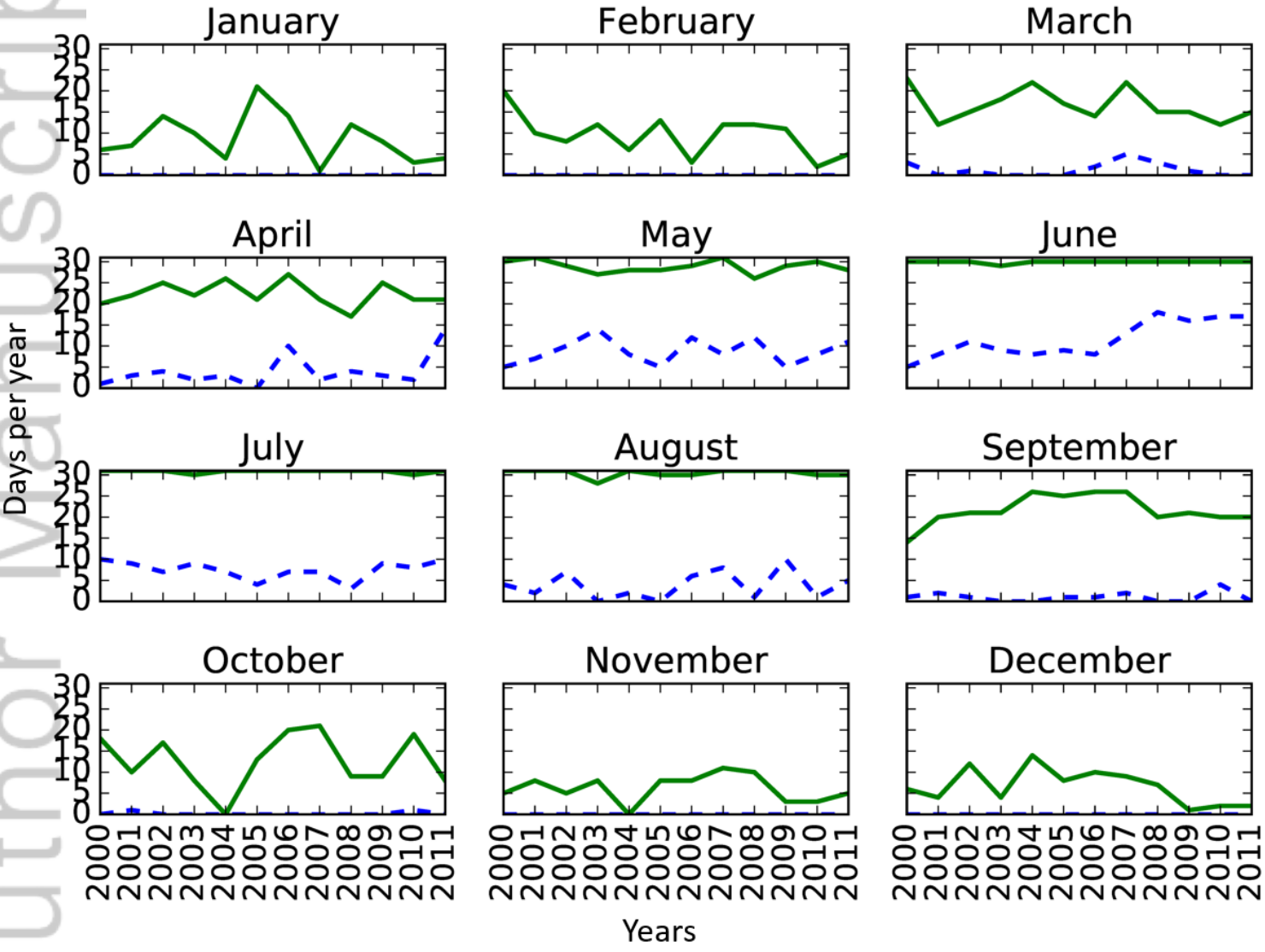
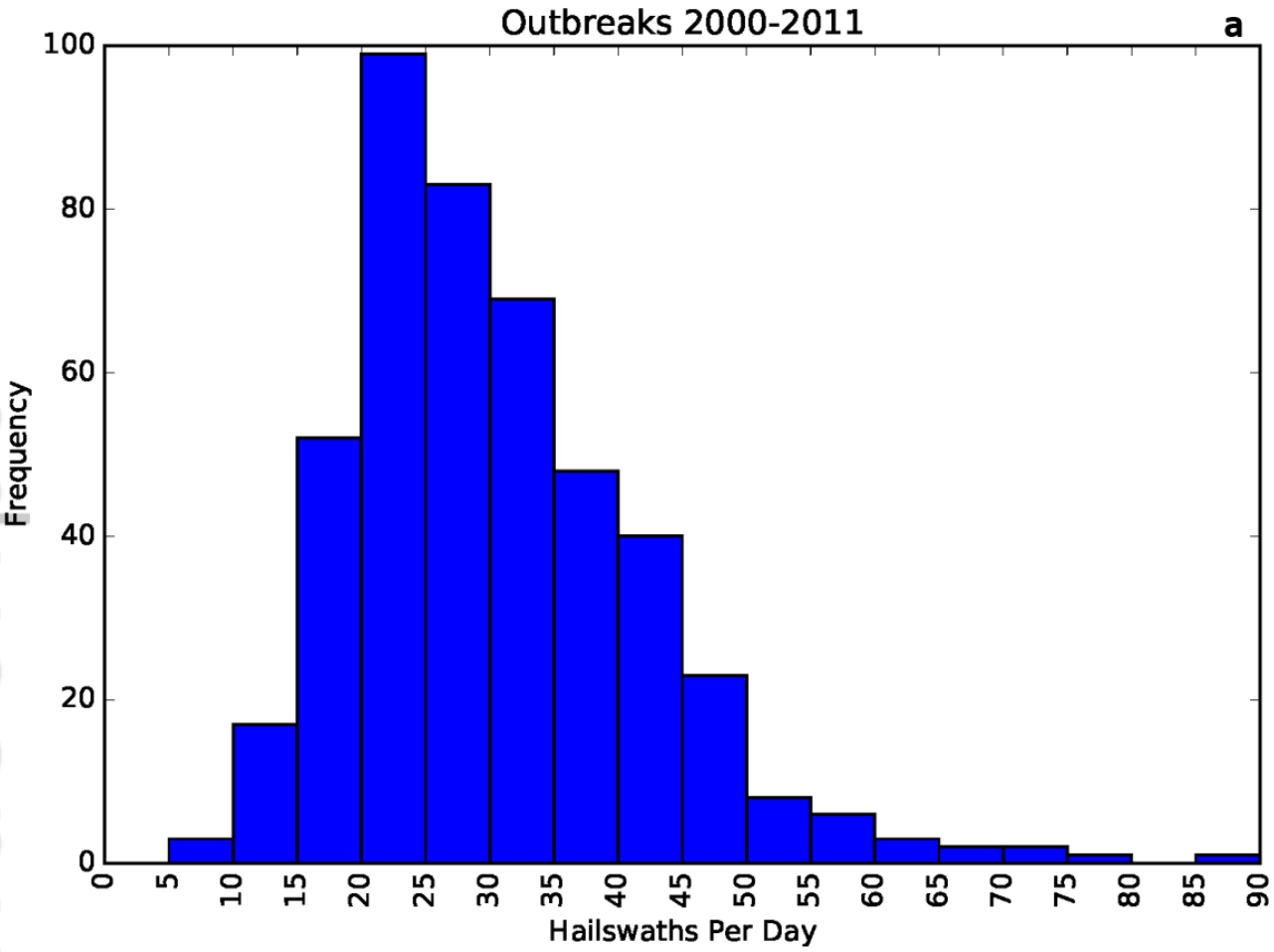


Figure10.tif



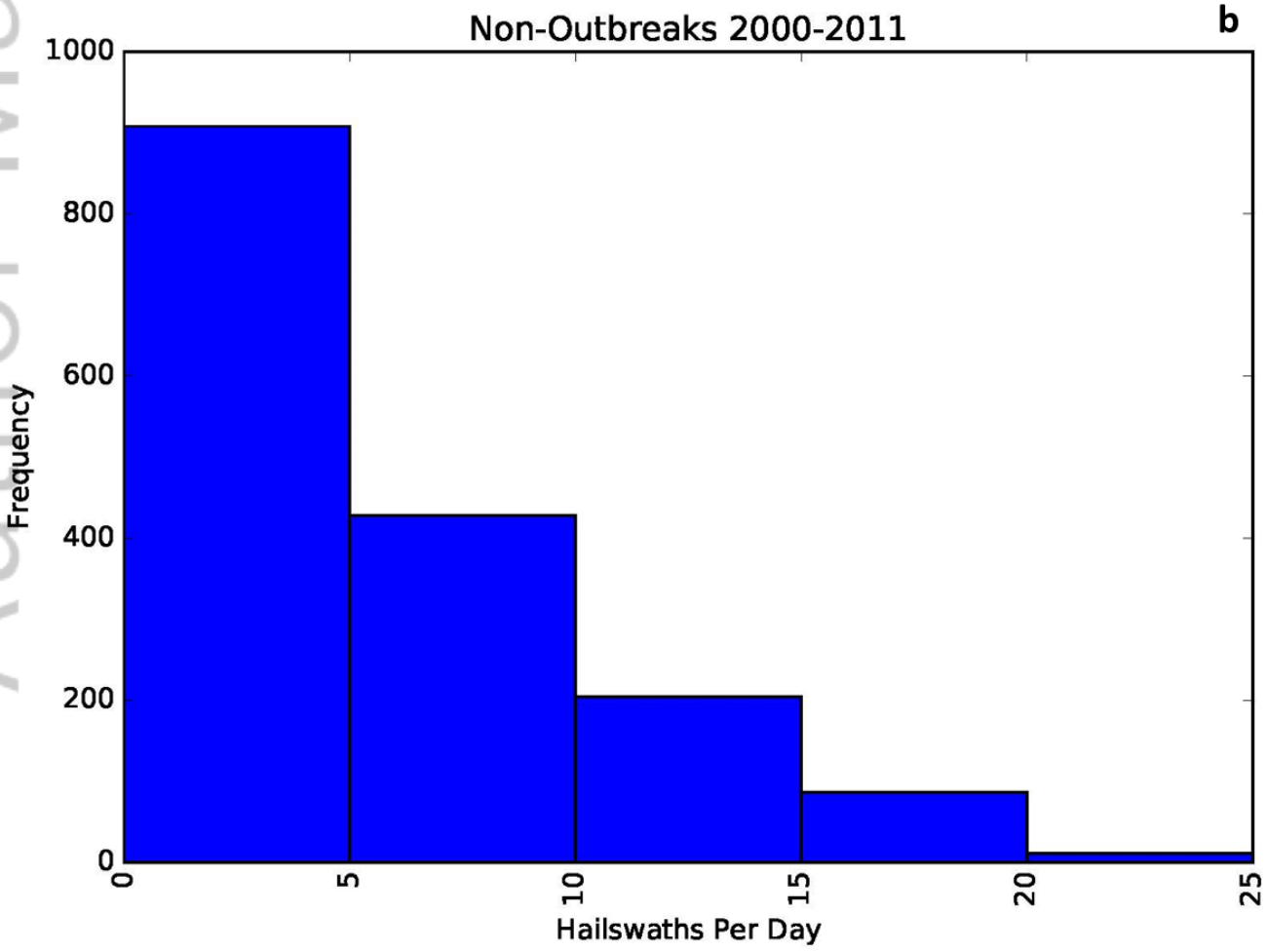


Figure11.tif

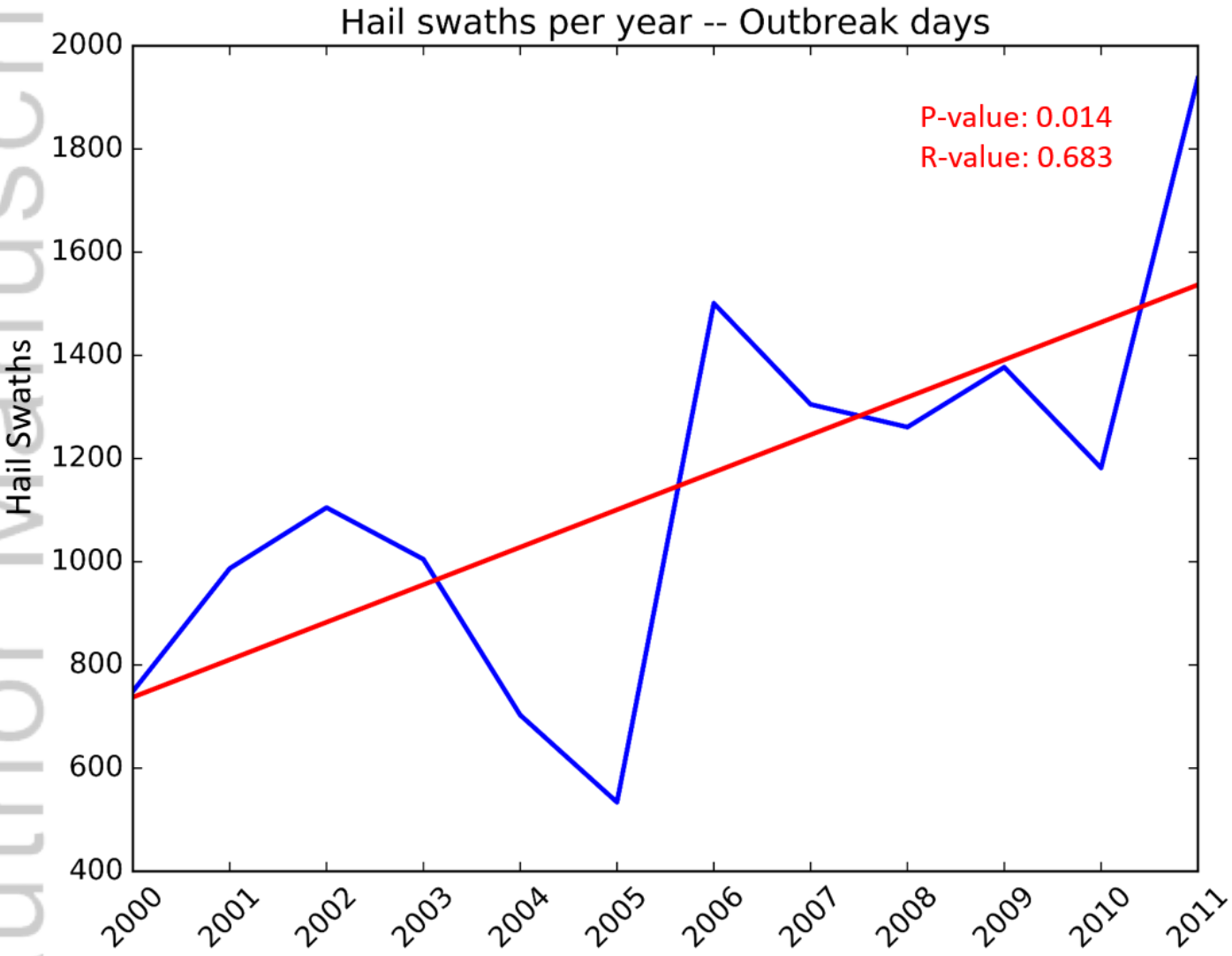


Figure12.tif

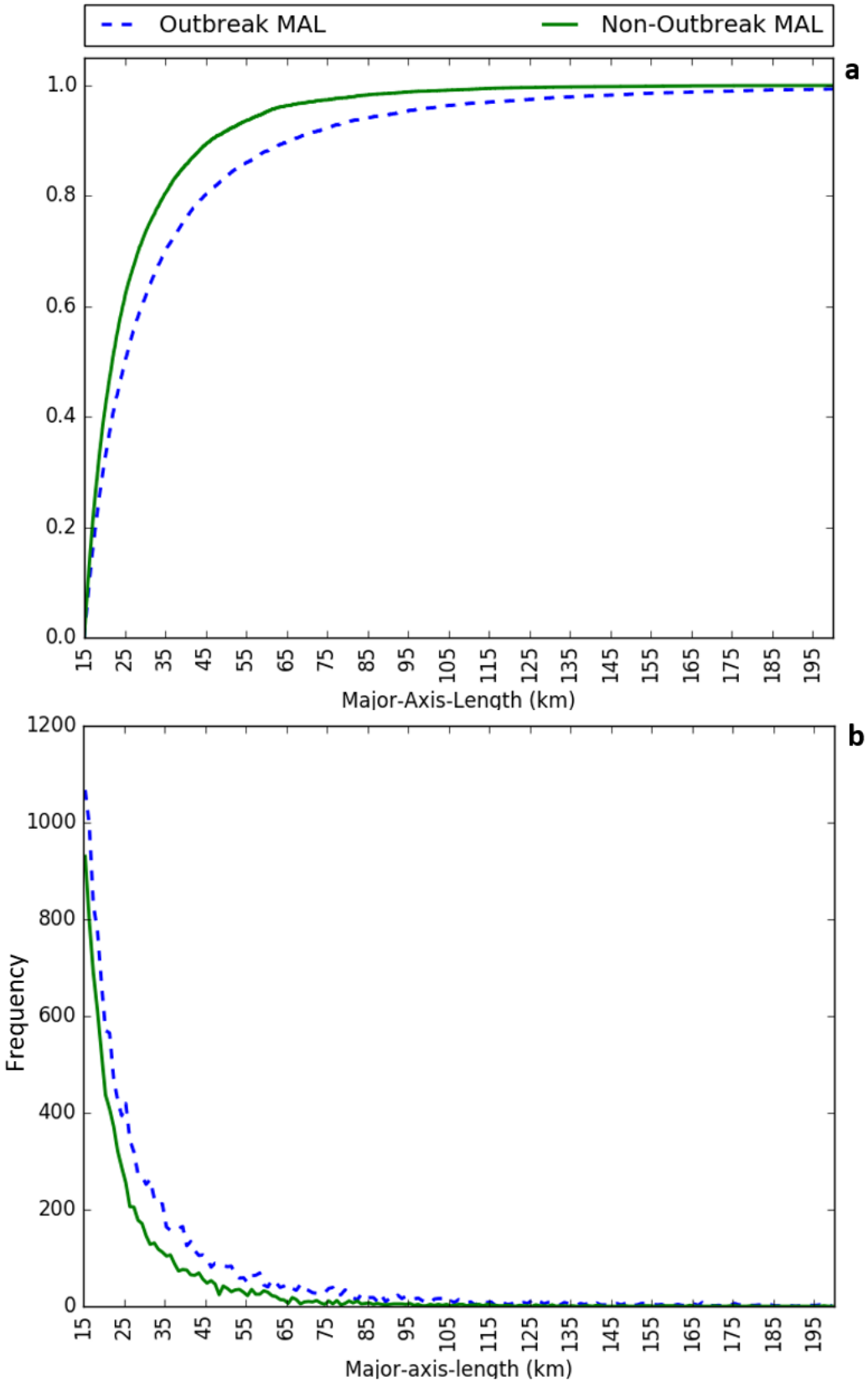


Figure13.tif

A radar-based study of severe hail outbreaks over the contiguous United States for 2000-2011

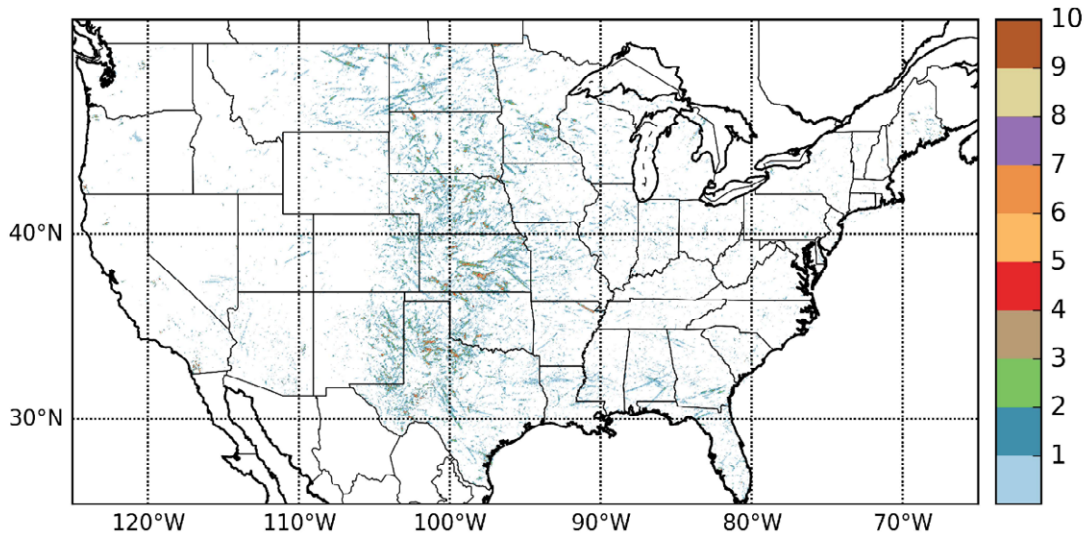
*Emily Elizabeth-Janssen Schlie

Donald J Wuebbles

Scott E Stevens

Robert J Trapp

Brian Jewett



A severe hail and severe hail outbreak day are defined in terms of the hail proxy Maximum Expected Size of Hail (MESH). This figure shows all severe hail MESH occurrences for 2005. We analyze characteristics and trends in severe hail, outbreaks, and hail swaths. We find a linear relationship between MESH and reports. Case studies are included to highlight the utility of MESH when studying outbreaks. We find that severe hail days decrease, while outbreak days increase from 2000-2011.

Date	MESH Area	Reports	MESH/Reports
04-10-2009	4947	387	12.8
06-15-2009	23644	386	61.3
03-12-2006	15537	594	26.2
07-22-2011	15458	93	166.2

Table 1: *Daily MESH area and official severe hail reports for 4 case studies. Top two show cases with similar number of reports and very different MESH area. Bottom two cases show similar MESH areas and very different total number of reports.*

Month	Days with Severe Hail Based on MESH area ≥ 100 grid points	Report based severe hail days	MESH based outbreak days (MESH area ≥ 6000 grid points)
January	8.7	5.1	0.0
February	9.5	10	0.0
March	16.7	20.8	1.3
April	22.3	26.7	4.0
May	28.8	30.5	8.8
June	29.9	30.0	11.6
July	30.8	30.9	7.5
August	30.4	30.8	3.8
September	21.7	24.3	1.0
October	13.8	17.8	0.2
November	6.7	8.1	0.0
December	6.6	4.8	0.0

Table 2: *Average annual number of days with severe hail based on MESH, severe hail based on reports, and severe hail outbreaks based on MESH (2000-2011). October and November of 2004 were omitted from the average due to corrupt data during that year. The R value for days with severe hail based on MESH versus reports is 0.974.*