

**Referee Comment #1:****Comment Line 167-168:**

Most papers that we used referenced in this paper specifically deal with European basins and they kept their basin areas smaller and applied size criteria. We wanted to get a better understanding Nationwide in the US and with basins varying drastically in characteristics and sizes we wanted to identify a common basin code to use. I did not come across any other paper that addressed Flood Frequency in the US using HUC's so we took it upon ourselves to choose two codes and work from there to identify which was more applicable. We chose HUC8 as our low end of our size range because as you can see in Figure 1, we are able to get approximately 25% coverage of the US, based on the criteria we applied to select sites, had we moved to a HUC10 we would have only be able to cover ~10% of the US which we felt was too low to accurately represent the US. With the HUC6 we were able to examine 85% coverage of the United States however we noticed the spacing of the sites might lead to more site events rather than events being attributed to the entire basin. With our goal being to identify events across a basin we wanted to eliminate the over estimation that you would see with more site events that should be aggregated to a single basin event.

**Comment Line: 201-204 and 208 – 209:**

We recognize the error that we made in by contradicting ourselves in our methodology. We are in the process of removing those sites with < 5 years of data. The outcome of this removal does not affect the overall frequency that drastically as only a few sites within each basin level was removed to account for this error. We will recalculate all of the necessary statistics to account for this change as well.

**Comment Line: 246 – 247**

We clarified the language during this portion to be more representative of what we wanted to state. Impacted area was changed to affected sites. Affected sites, is more applicable because we wanted to avoid running into errors where it might be interpreted as the flooding extent in terms of area. The affected sites is still an indication of the severity as you noted in the comment because we scale the discharges at those sites and divide by the number of sites in the basin.

**Comment Line: 266-267**

While this is intuitive where you would think that a larger the basin the higher the frequencies would make sense. Due to the fact that we do not select every single site in the basins, we needed to plot the relationship between events per year and catchment area as well as events per year and site count. From those figures, there is not a significant relationship between catchment area (square km) and frequency in both the HUC6 and HUC8; however we do need to represent a better comparison to further prove our point that the HUC8 provides a better representation of our method. The idea we painted with site count vs. frequency does show that there is a relationship between site count and frequency for both HUC6 and HUC8. The plot does show a stronger relationship for the HUC6; however this does not prove our hypothesis that the HUC8 is a better representation. We are currently working

on a plot that examines the distribution of the percentage of sites impacted during the events that we record. We are going to show that with our method we are seeing fewer localized events that should be aggregated to basin events when our method is applied to the HUC8 when compared to the HUC6.

**Comment Line: 280 – 281**

This language has been removed from the paper. It was missed during our earlier revision as it is not referenced in any other capacity during this paper.

**Comment Line: 285 – 292**

The shape of the CDF will be discussed in the revisions to this paper.

**Comment Line: 303 – 304**

The language of this statement will be clarified in the revisions to this paper. The sentence was meant to describe that when we see these locations being effected by natural phenomenon's, our definition of event duration as described by the paper does not factor in these events with a drastic event generation caused by an ice jam or prolonged duration from intense ground saturation. We wanted to point out that while it may seem that the duration of these events is longer than what you would expect, they are accurate because of our Q2 threshold that there may be remnants of flooding that are not severe but are still being considered and aggregated to the same basin wide events.

**Comment Line: 367**

The figure was updated to increase the size of the labels as well as place a blue outline around the selected HUC6 and HUC8 so that it would be easier to identify which basins the figure was referring to.

**Figure Comment:**

Figures were corrected to remove duplicates, increase label size and combined.

**Referee Comment #2:****Comment 1:**

This comment was made by referee #1 and we have corrected it in the manor listed below.

“We recognize the error that we made in by contradicting ourselves in our methodology. We are in the process of removing those sites with < 5 years of data. The outcome of this removal does not affect the overall frequency that drastically as only a few sites within each basin level was removed to account for this error. We will recalculate all of the necessary statistics to account for this change as well.”

**Comment 2:**

For this paper we wanted to give a basic understanding for our method of capturing flooding events within a basin on a nationwide approach. We selected the Q2 because of how previous literature states that it is the best approximation for “Bankfull Discharge.” We recognize that this may be slightly low at certain sites as noted from your comment above. For the purpose of this paper we wanted to test how our methodology would work with a standard metric at first. We feel that the best way to capture the understanding of our method is to run our algorithm on a simple threshold first to get a sense of its applicability and then within a response paper and further dissection of this method we will run a sensitivity analysis and re-run the algorithm on the same subset of events with thresholds of Q5, Q10 and potentially higher thresholds.

- This comment above is also a response to comment 1.

**Comment 3:**

We made the necessary changes to the figures that we felt should be combined as noted. We also removed certain figures in response to the previous referee’s comments to remove duplicate figures.

# **Defining and Analyzing the Frequency and Severity of Flood Events to Improve Risk Management from a Reinsurance Standpoint**

*Elliott P. Morrill<sup>1,2</sup>, Joseph F. Becker<sup>2</sup>*

<sup>1</sup>*Rosensteil School of Marine and Atmospheric Science at the University of Miami*

<sup>2</sup>*Guy Carpenter and Co. LLC*

*Correspondence emails:*

*Elliott.Morrill@guycarp.com*

*Joseph.F.Becker@guycarp.com*

1 **1.0 Abstract.**

2  
3 The National Flood Insurance Program debt has accelerated research into private flood  
4 insurance options. Offering this coverage begins with the ability to transfer the risk to the  
5 reinsurance market. Within the industry perils such as hurricanes and earthquakes have standard  
6 definitions but no such definition exists for flood. An event definition must examine the spatial  
7 and temporal aspects of the flood as well as the complexities of individual events. In this paper  
8 we were able to apply a data driven methodology to capture and aggregate flood peaks into  
9 independent events. Analyzing both the HUC8 and HUC6 a total of 7,932 HUC8 events and  
10 8,444 HUC6 events were recorded during the 15 water years used in our study. Each event was  
11 characterized by duration, magnitude and severity. Focusing on the HUC8, events were unevenly  
12 distributed nationally while severity was relatively evenly distributed. The goal for our study was  
13 to take a method and be able to apply it to basins of varying characteristics. This framework  
14 relied on the ability to analyze the individual processes related to each individual basin.

## 25 **2.0 Introduction:**

26           Throughout the world, flood events are one of the most destructive natural disasters.  
27 Floods occur for a variety of reasons, and risk factors such as total rainfall, soil types and land  
28 use can contribute to the complexity of events, in particular impacted area and event duration  
29 (Uhlemann 2010). Every year, major and minor floods contribute to economic and insured losses  
30 (Joyce 2014, FEMA). In the United States, the National Flood Insurance Program (NFIP) is the  
31 primary provider for residential flood insurance. Since its inception in 1968, the NFIP premiums  
32 have largely covered the amount paid out in losses (NFIP Act of 1968). However, the 2005  
33 Hurricane season, including Hurricane Katrina, which was the costliest storm in the program's  
34 history costing more than 16 Billion USD, pushed the NFIP into debt (Fig.C1). The NFIP debt  
35 was exacerbated by the significant property damage experienced during Superstorm Sandy in  
36 2012. Currently, the NFIP debt is estimated at \$24 Billion (Joyce 2014).

37           This extreme debt has accelerated research into a number of different private flood  
38 insurance options. One necessary issue to address before primary flood insurance can become a  
39 more standard offering is the ability to transfer risk to the reinsurance community. A challenge  
40 specific to flood is the complexity of individual events. Unlike the perils with an unambiguous  
41 event definition, such as hurricanes and earthquakes, there is no standard definition for a flood  
42 event, which can range in length from hours to months. The problem for flooding is not specific  
43 to the United States. In fact, reinsurers have offered flood risk transfer products in Europe and  
44 Asia for a number of years. For example, (re)insurers in Spain have provided flood insurance  
45 since 1971 (Barredo et al. 2012). Typically, reinsurance contracts define a flood event using an  
46 hour's clause ranging between 168 hours in the UK to 504 hours in Germany. Using the hour's  
47 clause insurance companies are able to aggregate claims during this period of time to limit

48 cumulative losses from multiple events (Munich Re. 2005). Defining events this way allows for  
49 providers to aggregate claims that can be associated with the same temporal event.

50         However, the hour's clause definition lacks the ability to discern between the shorter and  
51 longer events. Not all events can fit into a single defined time frame. If there are multiple short  
52 duration events occurring in quick succession then the claims from those events may be  
53 aggregated together. The hour's clause also lacks the ability to determine spatial aspects of each  
54 flood event. If events occur within the same window of time but in two different areas those  
55 flood are still attributed to one event. Aggregating these events limits the ability to understand  
56 the spatial extent based on impacted areas and the severity of each of the individual flood  
57 occurrences.

58         While research into flood event definitions is accelerating, it is not a novel topic.  
59 Research into event definitions has primarily focused on single site analysis (Bačová-Mitková &  
60 Onderka 2010, Mallakpour & Villarini 2016 and Kahana et al. 2002). However, as flood events  
61 are spatially complex, they often impact many locations limiting the use of single site definitions  
62 for reinsurance contract definitions. When events impact larger areas, multiple locations or entire  
63 basins, there is currently no method that can properly group flood peaks to the same event.

64         Public entities have compiled databases of flood occurrences to assist in frequency and  
65 severity analyses (NCDC). One goal of this type of analysis is to determine if floods are  
66 occurring more often and with increased severity due to climate change or other anthropogenic  
67 causes (Himmelsbach et al. 2015). Public databases are comprised of documentary sources and  
68 trained spotter observations (NCDC, EM-Dat, and DFO). The major downside of using this type  
69 of database to assist with reinsurance contracts is that they are based on subjective measures such  
70 as spotter definitions. Definitions follow a series of guidelines but varying flood characteristics

71 between regions can categorize flooding differently between these two regions. Variations in  
72 categorization have an impact on event durations and impacted areas. In addition to the  
73 definitions themselves, trained spotters respond to citizens reports of the peril. Depending on the  
74 area, what is considered abnormal flooding, in terms of standing water or bankfull discharge,  
75 may be reported in one area compared to another. For example an area such as Florida  
76 experiences significant precipitation year round which may contribute to minor flooding that is  
77 considered normal and thus not reported. However in an area like Los Angeles that similar minor  
78 flooding may be reported, which affects the frequencies of flooding in each area. Another source  
79 of flood occurrence information is using a documentary source, which involves examining media  
80 sources as well as government reports to comprise a set of occurrences across a state, country or  
81 globe (Himmelsbach et al. 2015 and Doocy et al. 2013). These sources rely heavily on the  
82 quality of the reporting, using the reports to assign severity and frequency estimates to cover an  
83 expansive region.

84 Relying on the quality of the reporting can lead to inconsistencies in what is reported and  
85 how it is reported. In a number of areas you can have two sources that can report statistics about  
86 an event that are drastically different. From those reports determining which is the most accurate  
87 becomes a challenge. Another issue with secondary sources is being able to define event  
88 duration. In many cases the reports cover the first instance of flood and damages associated but  
89 do not report the flooding on subsequent days defining the event duration. Spatially defining an  
90 event presents challenges. Not all events are reported equally across all areas. Secondary sources  
91 will primarily focus on the most severely impacted areas, but that provides a small picture of the  
92 entire event.



93 EM-DAT and National Climatic Data Center (NCDC) *Storm Data* databases are the two  
94 that are most commonly used datasets for this type of analysis. EM-DAT uses official records of  
95 areas affected, persons killed, disaster declarations issued and calls for international assistance  
96 made (EM-Dat, Doocy et al. 2013). The NCDC *Storm Data* database is a compiled set of  
97 observations from National Oceanic and Atmospheric Administration (NOAA) trained spotters.  
98 NCDC events are categorized by county and then separated by dates (Dobour and Noel 2005,  
99 Gaffin and Hotz 2000). EM-DAT catalogues events by year with summary statistics detailing  
100 frequency and overall event impacts (i.e. deaths and losses) from that year. Such summary  
101 statistics include injured, affected, total deaths and total damage. Both methods contain a number  
102 of different biases preventing use in reinsurance contracts including population biases, frequency  
103 biases and reporting biases. Due to the incomplete and often inconsistent reporting,  
104 implementing this method to formulate an event definition for reinsurance contracts presents a  
105 challenge. Despite their limitations, these datasets are useful first checks when developing a  
106 more robust method to define flood events as historical events can be compared to.

107 Many authors have shifted toward a data driven approach using the peaks over threshold  
108 analysis to examine changes in flood event frequency (Mallakpour and Villarini 2016, Bačová-  
109 Mitková & Onderka 2010), as well seasonality (Black and Werritty 1997). A data driven  
110 approach allows for the definition of an event to encompass a variety of basin characteristics.  
111 Authors choose a somewhat arbitrary threshold where if a peak observation exceeds the  
112 threshold, it is considered to be a peak over threshold (POT). A subsequent step for this method  
113 was to determine a metric for identifying independent peaks. Varying windows of time were  
114 used to identify the independence between the individual POT. Mallakpour and Villarini used an  
115 arbitrary window of 15 days, where any peak that occurs within this period is aggregated to a

116 single event. Black and Werritty determined their window by calculating the “time to rise” and  
117 identifying when the discharge dropped below  $2/3^{\text{rd}}$ s of the previous peak. Authors using these  
118 windows then looked at all individual peaks occurring within these windows to attribute them to  
119 the same event.

120 Site specific event identification is the base in developing a consistent method of event  
121 identification. However, our method will address the window of independence through an  
122 observational approach. Event independence should not be based on a standard window  
123 (Mallakpour and Villarini 2016). It must be based on how each site reacts to the flood waves.  
124 Implementing a concept similar to time to rise and a drop in discharge (Black and Werritty 1997)  
125 was the first of many steps taken toward resolving this. The window must cover the time before  
126 and after a peak, as previous peaks have an influence on succeeding peaks. Incorporating this  
127 into our definition will reflect the individuality of each site and the flexibility of our definition to  
128 cover a wider range of sites.

129 The primary goal of this research is to expand our definition to an entire basin or  
130 catchment area. These regionally impacting events are titled basin or “trans-basin” events (Nied  
131 et al. 2014, Uhlemann et al. 2010). Both papers used the POT method as well. Starting with a  
132 single site, individual events were identified (Uhlemann et al. 2010) and then all mutually  
133 dependent events were identified from a moving temporal window. The window defined from  
134 previous literature provides a solid structure but categorizes catchments and basins into an all-  
135 encompassing time frame. A more basin specific time frame is measurable and would not  
136 underestimate the smaller basins or overestimate the larger basins.

137 This paper seeks to define events through a data driven approach aimed at accounting for  
138 the individuality of flood waves and the basins they impact. Our main goal is to develop a

139 consistent definition in order to examine how frequency and severity vary regionally. Looking at  
140 frequency regionally provided us with a clearer picture of the specific areas that were more at  
141 risk for flooding. Severity allowed us to look at how areas with similar frequencies were  
142 experiencing events in terms of impacted areas and overall magnitude. Severity will factor into  
143 future implementation of risk mitigating factors that can look at two areas and determine the  
144 steps needed to protect a certain area. It also allowed us to determine if our method is  
145 representing more local or extreme flooding across the various basins.

146         Methods implementing the hour's clause or standard event windows lack the ability to  
147 interpret how each individual flood wave progresses. Understanding the individuality of the  
148 flood is the basis for how our method will tackle a standard event definition. This paper will be  
149 structured as follows: Section 2 will cover the data availability as well as the data selection  
150 process along with which tools were used to analyze the data. The concepts that feed into our  
151 method as well as our method itself will be discussed in Section 3. Section 4 will provide the  
152 results of the analysis from our methodology with comparisons to methodologies exhibited in  
153 previous research. Section 5 will provide the discussion and concluding remarks regarding our  
154 results within this study.

## 155 **2.0 Site Selection:**

156         This research focuses on expanding the definition of a flood event from an individual site  
157 to river basin. As this research focuses on the United States, USGS daily flow gauges stations  
158 were used to identify individual sites and USGS Hydrological Unit Codes (HUC) were used to  
159 define river basins. River basins can be defined in a number of ways and determining the  
160 appropriate size can be a non-trivial task. For use in reinsurance contracts, river basin should be  
161 defined in such a way that flooding events within a portion of the basin show a correlation to

162 events in other portions. A river basin needs to be defined in such a way that we can see how  
163 flood waves impact the basin and not individual sections of that basin. The USGS HUC codes  
164 follow the Pfafstetter Coding System meaning that each unit code is delineated in a hierarchical  
165 fashion. Drainage areas are defined on a continental scale and then divided and subdivided into 6  
166 levels. Each level is associated with number of digits corresponding to size. Digits range from 2  
167 – 12, largest to smallest (USGS), with the 8/6 digit HUC's being used. A majority of the papers  
168 that we referenced in this study have dealt with European or Asian basin definitions and were  
169 focused on one or two basins within a finite area. With our broad scope of study, we needed to  
170 look at basins across a variety of characteristics so a common basin code was needed for  
171 comparisons of frequency. Other research of flood frequency did not yield any references to the  
172 HUC basin codes so as authors we developed our own criteria that we felt best represented the  
173 size of the basins most applicable for our methodology. Our decision to use these two size  
174 HUC's relied on looking for a basin size that allowed us to observe how the events would  
175 aggregate to a basin level event rather than being identified as two separate events. We wanted  
176 our dataset to contain the largest percentage of HUC's possible after our site selection criteria to  
177 get a better nationwide picture of how our method observed basin wide flood events. With the  
178 HUC8 we were able to get approximately 20% of coverage across the United States with a basin  
179 contained all 20 HUC2's (Fig.1). With any HUC size below the HUC8 such as the HUC10 we  
180 were left with a much lower coverage percentage, roughly less than 10% for the HUC10, which  
181 would not accurately represent the methodology across the country. When we look at the upper  
182 end of our HUC size for the HUC6, we when look at how frequency compares with site count  
183 above the HUC6 we saw that frequencies were heavily affected by site count. From these two  
184 factors we felt that the HUC8 and HUC6 were the most applicable basin sizes. Daily mean

185 discharge as well as Annual peak streamflow was used for all sites, which provided data for  
186 those parameters.

187 From all available HUC's, sites and basins were selected based on a number of selection  
188 criteria. The first criteria removed sites with less than 5 years of daily discharge data. The second  
189 criteria required sites to occur along natural rivers and streams; gauges impacted by reservoirs  
190 and other impediments to natural flow were excluded. Following site removal, HUC's with less  
191 than 5 sites were excluded. Finally, HUC's were required to have at least 3 sites that overlapped  
192 with 70% of the data during each individual year that was examined. Due to the nature of our  
193 method seeking to aggregate peaks from multiple sites, the sites needed to overlap or else that  
194 method would be looking primarily at individual site events instead of the basin events. Of the  
195 2,300 HUC8's and 387 HUC6's available, 462 HUC8's and 276 HUC6's were used (Fig.1) with  
196 a total of 3,121 and 4,919 gauge stations within the HUC8 and HUC6 respectively. Both HUC  
197 sizes were analyzed for initial frequencies and the most applicable HUC was chosen for  
198 subsequent analyses.

### 199 **3.0 Methodology:**

200 Daily discharge data from 8,084 river gauge stations was obtained from the USGS  
201 ([http://nwis.waterdata.usgs.gov/nwis/dv/?referred\\_module=sw](http://nwis.waterdata.usgs.gov/nwis/dv/?referred_module=sw)). A study period of 15 water  
202 years between 2000 and 2015 was selected for this analysis. Initial attempts to expand the period  
203 of analysis severely reduced the number of basins that met the criteria for analysis. The peak  
204 over threshold method outlined in Uhlemann et al. (2010) was conducted on all basins that fit the  
205 criteria for analysis. The peak over threshold method consists of identifying individual  
206 observations over a specified threshold within a particular time window. The procedure was split  
207 into 4 major steps: (1) identifying peaks occurring at each site within each basin and the

208 subsequent peaks over threshold; (2) applying a window of independence at each site to  
209 determine independent site specific events; (3) compiling all independent site specific events and  
210 applying a secondary window of independence to determine independent basin specific events;  
211 (4) applying multiple characteristics to determine a severity score to compare differing events  
212 from one another.

213         The first step involved selecting a minimum threshold. The median of annual maximums  
214 was chosen as the threshold in which a flood peak must exceed. The median of annual  
215 maximums was chosen because it corresponds to the 2-year quantile, or Q2. Uhlemann et al.  
216 (2010) states that the “Q2 is a rough estimation for bankfull discharge on naturally occurring  
217 streams.” For sites with at least 5 years of annual peak streamflow data, their Q2 was calculated  
218 by taking the median across the entire time series. As peak discharges are determined by  
219 instantaneous measurements, small catchments can exhibit extreme values, which are rarely  
220 observed in the daily record. The extreme values may lead to a minimum threshold that may not  
221 be a representative measurement of flooding for that catchment area. The discharge at each of the  
222 peaks recorded, were then compared to their respective sites Q2 value to determine all of the  
223 peaks over threshold.

224         The next step in identifying site specific events is to determine a time criteria that defines  
225 independent site events. Two metrics were calculated for all peaks over threshold to determine  
226 the duration of each event: base to peak (BtoP) and peak to base (PtoB). Base to peak is the time  
227 it takes for the discharge to reach the peak after it has crossed the minimum threshold. Peak to  
228 base is the amount of time it takes for the discharge to return to the minimum threshold  
229 following a peak (Fig.2a). In the case where there are multiple peaks before the discharge returns  
230 to base, the peak was selected as the observation that experienced the maximum discharge. Each

231 peak over threshold has a unique BtoP and PtoB that could have a significant range. To  
232 standardize the windows of independence for each site the median of both metrics was calculated  
233 and then the peaks start and end times were recalculated. Our window of time was aimed at  
234 eliminating the extreme events on either end of the temporal distribution to determine a window  
235 that reflected the time it would take for a flood wave progress through a site.

236         After the windows were recalculated, combining peaks with overlapping or consecutive  
237 windows into a single site specific peak consolidated peaks. All peaks over thresholds with  
238 windows that did not overlap were treated as independent events. Each event was characterized  
239 by, site number, start time, peak time, end time and peak discharge. For the peaks, which  
240 overlapped, the start time was defined as the earliest start day and end time was the latest end  
241 date. The peak discharge from each event was then scaled by the Q2 at each site. Scaling each  
242 peak discharge reduced the impact of catchment size when comparing magnitude of discharge  
243 and made the different sites comparable.

244         A similar methodology of consolidating overlapping observations was applied to define  
245 basin specific events from the site specific events (Fig.2b). The basin specific events used the  
246 start and end time of each site specific events that occurred within the basin. If the windows of  
247 time between the start and end of the site specific events overlapped or were consecutive (i.e.  
248 occurred within 1 day of another peak), then these events comprised one basin specific event.  
249 The start of the event was the earliest start time recorded at any site and the end of the event was  
250 the final end time recorded. Each event was defined by start time, end time, peak time, and peak  
251 discharge for all events from the desired HUC's.

252         The final step involved determining a severity score for each basin event. Defining  
253 severity allowed us to compare areas of like frequency. From these we were able to see the

254 certain areas that are more vulnerable during flooding. Severity scores in future analyses will  
255 also factor into pricing of reinsurance contracts. Severity of each event was designed to include  
256 elements of the spatial extent as well as the magnitude of the flooding experienced in the basin  
257 by the affected sites during each event. The severity score represents a number between 0 and  
258 infinity where the high value indicates a more severe event. The affected sites were defined as  
259 the number of sites within the desired HUC, which recorded a peak over threshold during the  
260 event. Total discharge was the sum of the discharges, scaled by their corresponding minimum  
261 threshold, observed at all the affected sites. Severity was calculated by taking the sum of all  
262 scaled discharges and dividing by the total number of sites within the basin, *EQ.A1*. If a site was  
263 impacted more than once during a basin event, the maximum-scaled discharge was selected to  
264 calculate the severity score. Scores less than one are expected when looking at the minimum  
265 threshold as it represents small scale and localized flooding, in terms of discharge and the  
266 percentage of sites it may impact within the individual HUC.

267 From the analyses, we compared the HUC6 and the HUC8 to determine which size basin  
268 was more appropriate for our method. For each HUC aggregation, frequency, event duration and  
269 severity distributions were examined. Two comparisons were made to the NCDC *Storm Data*.  
270 The first method looks at all reports of flooding and aggregates them by county. The second  
271 method used a standard 13-day independence window, 3 days pre-peak and 10 days post-peak  
272 (Uhlemann et al. 2010). A standard window was used because the NCDC observations are  
273 unable to provide a site specific window of independence.

#### 274 **4.0 Results:**

275 A total of 7,932 and 8,444 events were calculated for basins defined by the HUC8 and  
276 HUC6 respectively. Table B1 provides the frequency summary statistics for both the HUC8 and



277 HUC6 basins. Comparing the frequency distribution of events between the two selected basins  
278 sizes suggests that frequencies within basins defined by the HUC6 are higher than frequencies  
279 defined by the HUC8 (Fig.3 & Fig.4). We can see that from Figure 3, the frequencies in each  
280 HUC8 are typically lower than the frequencies found each HUC6. This is highlighted in Figure  
281 4, where we focus on 6 HUC8's that make up 1 HUC6 (Outlined in Blue). From here the  
282 individual basins in the HUC8 indicate a lower basin level frequency than at the HUC6. This  
283 comparison is important because the aim of this paper is to define events at a basin level by  
284 aggregating individual events into basin wide events. To explore this concept more we wanted to  
285 look at the impacted sites during the events compared to the total number of sites within the  
286 basin to get a sense of how many events are being determined as local when they should be  
287 aggregated. While there will be a number of small local floods that this methodology captures,  
288 we looked at this to provide us with an indication of whether the HUC8 is too small of a basin to  
289 use or the HUC6 is too large.

290 We looked at the distribution of the percentage of impacted sites by event for each HUC  
291 (Fig.5). We took each event within our catalog and identified how many sites were impacted.  
292 The percent impacted was calculated by taking the number of sites impacted and dividing by the  
293 total number of sites within the basin. For the events within the HUC8 on average 36% of the  
294 sites were impacted compared to 21% for the HUC6. When you look at the CDF of the events of  
295 HUC6 and HUC8 (Fig.5), we can clearly see that the HUC6 events impact a fewer percentage of  
296 sites. While HUC6 does have more sites, due to our methods intended aggregation of events we  
297 would expect a similar percentage of sites impacted between the two. However, because the  
298 HUC6 is showing a lower percentage of sites impacted during the events in their catalog this is  
299 an indication that the HUC6 does not aggregate individual events as well as the HUC8. 80% of

300 the events within the HUC6 have % impacted <40% compared to the the HUC8 where ~50% of  
301 their events are impacting 50% of the sites. Due to the size of the HUC6, the basin is being  
302 segmented during our method and is not capturing events that should be attributed to the same event.  
303 The segmentation of the events within the basin will lead to an overstating of the frequencies.  
304 Overall, the HUC8 is showing a higher percentage of events in the higher percentages of sites  
305 impacted meaning that our method is aggregating individual events into basin events at this basin  
306 size. From both the CDF and the average we have concluded that the HUC8 is a more applicable  
307 basin size due to its ability to aggregate the events within the basin rather than segmenting them.

308         Nationally, the median frequency of events HUC8 basins was 1.00 events per year while  
309 the mean was 1.14 events per year (Fig.6). This frequency varied regionally with some areas  
310 experiencing higher frequencies (Fig.1 Left Panel). Notable population centers that experience  
311 elevated frequencies include the Upper Midwest (south of Lake Michigan), Southern California  
312 and Southern Florida. For the HUC6 basins, the median frequency of events was 1.87 events per  
313 year with a mean of 2.03 events per year (Fig.7). Similarly to the HUC8 basins, the frequencies  
314 varied regionally with some areas of elevated frequencies (Fig.1 Right Panel).

315         To investigate how event duration varies nationally, we calculate the mean event duration  
316 for each basin. Nationally, the mean event duration ranged from two to 79 days for the basins  
317 defined by the HUC8 and two to 73 days for the basins defined by the HUC6. The mean event  
318 duration for 95% of HUC8 and HUC6 basins is less than 14 and 17 days respectively (Fig.8 and  
319 Fig.9). The minimum event duration was two days and was observed at 336 HUC8's and 227  
320 HUC6's. The maximum event duration for HUC8's was 232 days and occurred in the 10160003  
321 basin. For HUC6 basins that maximum event duration was 237 days occurring in the 101600  
322 basin. When we look at the shape of both curves, we can see that there is a higher percentage of  
323 HUC8's that have shorter mean and maximum durations, as the curves approach the lower event

324 durations more rapidly leading to a steep curve when compared to the curves for the HUC6.  
325 However, when we look at the minimum duration, a larger percentage of the HUC6's have a  
326 minimum duration of 2 days when compared to the HUC8 which is an indication that there is a  
327 larger number of events that are impacting only one site. While both the HUC6 and the HUC8  
328 taper of towards the higher event durations, there is a lower percentage of the HUC8's that have  
329 event durations greater than 20 days. With those two factors we can see that durations within  
330 each HUC6 have a wider range than those compared to the HUC8.

331 Figure 10 represents two sites that reflect longer recession periods following their peaks.  
332 With a data driven approach identifying the generation and recession of the events, certain  
333 extreme events may show increased event durations. The extreme durations are a reflection of  
334 the minimum threshold as well as the hydrological processes at hand. Looking at the two sites,  
335 the left is located in South Dakota and the right is located in Florida; both of the extreme events  
336 that are observed have certain factors that impacted their recessions. The site in South Dakota  
337 experienced an event that was impacted by the melting of an ice jam represented by the quick  
338 generation. Following the melt there was a significant rain event as well as a release of water  
339 from a dam further upstream. The site on the right is located on a natural tourist spring. These  
340 springs contain a significant amount of ground water. Following an intense rain event the  
341 buildup of water caused the increased recession. When we define an events' duration as the first  
342 occurrence of discharge above the Q2 to the final occurrence of discharge below the Q2, if our  
343 site is impacted by a natural occurrence, events will reflect longer than expected durations. These  
344 durations are longer than we would expect and further analysis will be conducted to examine  
345 changes to the minimum threshold to examine the influence of these natural processes. While a

346 majority of the durations reflect reasonable time frames for flooding events that exceed the Q2 it  
347 is important to note that the method might not be appropriate for all streams.

348         When looking at the distribution of severity scores there is a slight skew towards the  
349 extreme events. Severity scores ranged from the least severe, 0.032 to the most severe, 26.9  
350 (Fig.11) with a median severity score of 0.32 and a mean of 0.57. While the range in severity  
351 scores is quite large, a majority of the events received a score less than 1. Regionally the severity  
352 scores are generally distributed evenly throughout the country (Fig.12). There appear to be  
353 pockets of higher severities but across the country there does not appear to be a pattern within  
354 the regional distribution. While it is evenly distributed regionally, within the regions we can see  
355 the wide range in severity that was observed in the distribution of frequency.

356         Finally, comparisons were made to other methodologies applied to the same dataset as  
357 well as other publically accessible datasets. The first comparison examined a method used by  
358 FEMA to estimate floods using NCDC Storm Events Database (Fig.13). The distribution of  
359 events was broken down into total event frequency by county ranging from one to 4,114. While  
360 the trained spotters follow guidelines in identifying events, the method lacks a way to group  
361 events. The inability to group events that would otherwise be considered a single event, leads to  
362 an overestimation of events. This overestimation is evident when it is noted that the maximum  
363 frequency of events for a specific county was 4,114.

364         The final comparison was made to the NCDC applying a 13-day standard window. While  
365 the NCDC map provides a more complete national coverage two patterns occur (Fig14). Within  
366 the 5-boxed areas, either the NCDC frequency is far greater or the daily discharge frequency was  
367 far greater. For example, in Florida, we see frequency range from 6 to 25 events for NCDC  
368 observations but events observed through daily discharge range from 26 to 45. The opposite

369 occurs in Missouri with NCDC estimates ranging from 16 to 85 events with events observed  
370 through daily discharge ranging from 6 to 15.

371 From these estimates there is no obvious reason for the discrepancies in frequencies but  
372 we can speculate. For example Florida experiences significantly fewer events using NCDC data  
373 than the daily discharge data. A possible explanation could be how trained spotters define events.  
374 An area in Florida may experience a peak over the threshold triggering our event definition, yet  
375 that peak may not be recorded as an NCDC observation based on the spotters perspective.  
376 Another reason could be due to the fact that these trained spotters respond to citizen's reports  
377 and, due to the frequency of flooding in an area like Florida, the citizen may not call and the  
378 peak may not be recorded.

379 However a similar thought process can be applied to our threshold selection. As stated  
380 the minimum threshold was selected as a representation of bankfull discharge. While this  
381 assumption was the basis for our method, in certain areas it is conceivable that the threshold may  
382 be lower than bankfull discharge which could possibly lead to an over estimation of flooding  
383 events in certain areas. There is no certain explanation for the discrepancies in the results. With  
384 no certain explanation for the results from this comparison, the assumptions that define the  
385 compared methodologies will be explored in future analyses.

## 386 **5.0 Discussion and Conclusions:**

387 This study was able to provide a data driven approach in attempts to solve the issues of  
388 inconsistent event definitions within the (re)insurance industry. We derived a methodology based  
389 on a peak over threshold analysis that was able to capture and aggregate multiple occurrences of  
390 flooding at various locations. Using physical assumptions, our minimum threshold and window  
391 of independence were able to capture each individual sites reaction to passing flood waves. An

392 approach identifying windows based on the impacted site allows for each site to represent their  
393 individual characteristics of flooding rather than applying standard metrics throughout. Each  
394 event was defined through their duration, impacted area and magnitude. The development of a  
395 severity index examines overall impacted areas as well as individual flood magnitudes.

396         Analyses were conducted on both HUC8 and HUC6 to determine which size of  
397 Hydrological Unit Code was more applicable for further analysis. 7,932 HUC8 and 8,444 HUC6  
398 events were identified during our study. Understanding the applicability of different basin sizes  
399 is important because it aids in our main goal of applying a consistent definition to reinsurance  
400 contracts. From our definition our goal was to understand the frequency that represents an entire  
401 basin or area. We also hope to use the definition to define a parametric trigger or an alternative  
402 form of defining the event. All of this is possible when we know what basin size is the most  
403 applicable. The HUC8 was chosen as a more applicable basin size as it was a better  
404 representation of site interaction during flooding events.

405         Nationally, there are areas with large discrepancies between the HUC6 and HUC 8  
406 frequencies. One explanation of this discrepancy is represented by HUC6: 071200 (Fig. 4). The  
407 area of this HUC6 is 28,309.78km<sup>2</sup> and contains 6 HUC8s. The annual frequency of events of  
408 the HUC8 ranges between 1 and 2.33, while the HUC6 produces 5 events per year. Although it is  
409 expected that the larger basin will have a slightly higher frequency due to some events occurring  
410 in one part of the basin and not impacting the other, a more than doubling of events per year  
411 indicates that a large number of events do not interact with other sites in the basin. This lack of  
412 interaction is inconsistent with the goal of this research to identify basinwide event frequencies.  
413 The inconsistencies and lack of interaction are represented by the relationship between site count  
414 on frequency (Fig.5).

415 We found that HUC8 frequencies are relatively normally distributed but are unevenly  
416 distributed regionally. For all HUC8's a median of 15 events (1 event per year) and mean of  
417 17.21 events (1.14 events per year) were recorded. In a number of areas there were pockets of  
418 elevated frequencies. Durations for all events ranged from 2 – 232 days with a mean duration of  
419 6.34 days. The wide range of event durations prompts further investigation into events with  
420 durations in the positive tail of the distribution. For example, we considered two HUC8's, one in  
421 South Dakota (10160003) and another in Florida (03100207), that are impacted by natural events  
422 leading to longer durations. Some sites within these two basins were affected by ice jams as well  
423 as natural springs, which have contributed to significant recessions of their events. While these  
424 events are natural, the resulting event durations should prompt examination into the selection of  
425 thresholds for the sites, as an assumption of bankfull discharge might be slightly lower than a  
426 threshold that produces flooding.

427 Severity scores calculated for all events in the dataset showed a slight skew toward the  
428 more extreme events. The smaller and local events are represented by the median of 0.32 and  
429 mean of 0.57, as we can expect events slightly above the threshold to not necessarily affect all  
430 the sites in the basin, producing a score less than 1. Regionally severity is relatively evenly  
431 distributed nationally.

432 With a data driven approach to our methodology, a focus on the individual site  
433 parameters shifts the focus from generalities about events to site specific understanding leading  
434 to an applicable method regionally. A fundamental aspect of this research is to understand spatial  
435 extent of flooding and we were able to expand from single gauge stations to entire basins. The  
436 data driven approach allowed us to apply the methodology to a number of basins with varying  
437 characteristics. The final advantage to our method is that when looking at flood severity we do

438 not look at exclusively magnitude but the addition of spatial extent adds an element to  
439 differences in severity regionally.

440 While there are a number of advantages that come from this method, relying on public  
441 data have revealed drawbacks in its application. Being a data driven method limits our ability to  
442 estimate frequencies in areas that do not have data. Across all USGS gauges there is no  
443 uniformity in data availability for number of years or number of stations within a basin. Through  
444 our site selection process we were only able to use 20% of all available HUC8's, which limits  
445 national coverage in our estimates.

446 The minimum threshold for flooding is based on the assumption that it is a representation  
447 of bankfull discharge; in certain areas this may not be accurate. Riverbanks are not uniform so  
448 how bankfull discharge is recognized at each site is dependent on that location, which may lead  
449 to underestimation or overestimation of flood stage at that site. The final drawback we observed  
450 was that when taking the median of the BtoP and PtoB slight variations in the event window  
451 occurred on the more extreme events. Instead of median other statistics will be tested to  
452 determine the most applicable way to represent the basin flood generation and recession.

453 For further research a comparative analysis will be conducted altering the threshold to  
454 examine how that might affect frequency as well as severity. Increasing the time frame will also  
455 provide insight as to whether or not this 15-year period is representative of the entire time frame  
456 of data or if we see a significant increase in events during certain subsections. Seasonality tests  
457 will be run to observe areas more frequent and more severe times of year which may also  
458 provide insight for risk managers. The final test that will need to be conducted is a sensitivity  
459 analysis on the threshold selected to prove which threshold is the most reasonable for an analysis  
460 such as this.



461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506

507 **Code Availability:**

508

509 All calculation and download scripts have been included in the supplemental folder. All  
510 scripts were written using R-Studio.

511

512 **Data Availability:**

513

514 All data is publically available from the NCDC Storm events database as well as the USGS  
515 stream gauge data sites. A list of sites and a list of the years used will be included as well as the  
516 compiled file of the data, added to the supplemental files.

517 <ftp://ftp.ncdc.noaa.gov/pub/data/swdi/stormevents/csvfiles/>

518 <https://waterdata.usgs.gov/nwis/uv>

519

520 **Team List:**

521

522 Elliott P. Morrill

523 Joseph F. Becker

524

525 **Author Contribution:**

526

527 E. Morrill and J. Becker designed the methodology. E. Morrill wrote and executed code to  
528 carry out the methodology. E. Morrill performed the manuscript with help from J. Becker.

529

530 **Competing Interests:**

531

532 The authors declare that they have no conflicts of interest.

533

534 **Disclaimer:**

535

536 The opinions expressed by authors contributing to this journal do not necessarily reflect the  
537 opinions of the Hydrology and Earth System Sciences Journal or the institutions with which the  
538 authors are affiliated. The data and code used within this research is a property of Guy Carpenter  
539 and Co. LLC.

540

541 **Acknowledgments:**

542

543 I would like to acknowledge the support of Guy Carpenter LLC and the Nat/Geo group  
544 within the Analytics Department. I would also like to thank my advisors from the University of  
545 Miami, Igor Kamenkovich, David Letson, Roni Avissar for supporting me during my time at The  
546 University of Miami as well as my time at Guy Carpenter and Company LLC.

547

548 **Appendices:**

549

550 **Appendix A.**

551

$$\text{Severity} = \frac{\sum Q(i)_{\text{scaled}}}{\# \text{ of Sites (HUC)}}$$

552

553 *EQ.A1. Severity Score*

554

**Appendix B.**

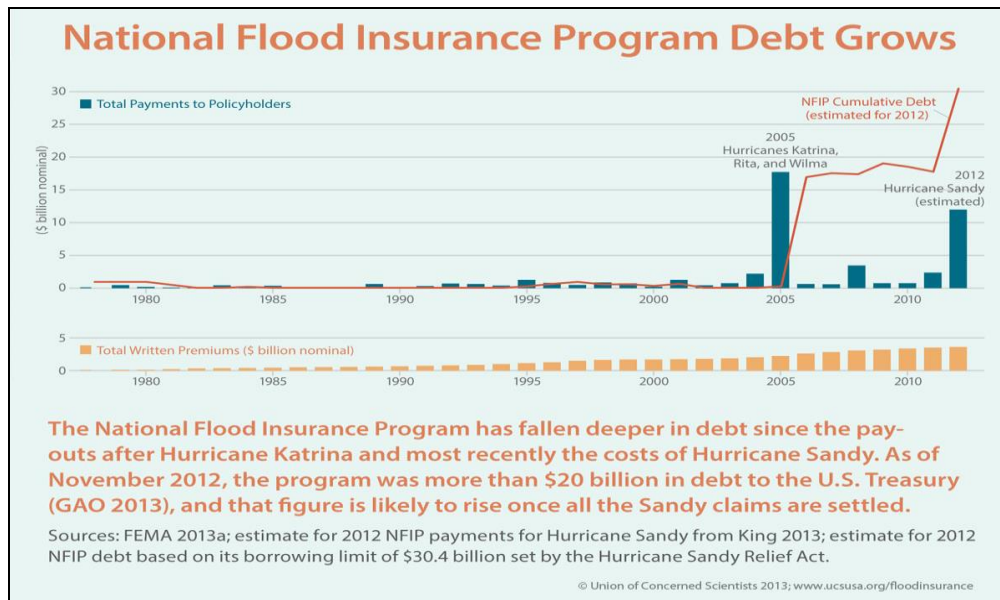
| HUC | Total HUCS | Selected HUCS | Minimum Freq. | 1st Quantile | Median Freq. | Mean Freq. | 3rd Quantile | Maximum Freq. |
|-----|------------|---------------|---------------|--------------|--------------|------------|--------------|---------------|
| 08  | 2300       | 462           | 0             | 10           | 15           | 17.17      | 21           | 63            |
| 06  | 387        | 276           | 0             | 19           | 27           | 30.59      | 38           | 145           |

555

*Table.B1. HUC8 and HUC6 Frequency Summary Statistics*

556

**Appendix C**



557

*Fig.C1 NFIP Cumulative Debt, Total Payments and Total Premiums, 1978-2012*

558

559

560

561

562

563

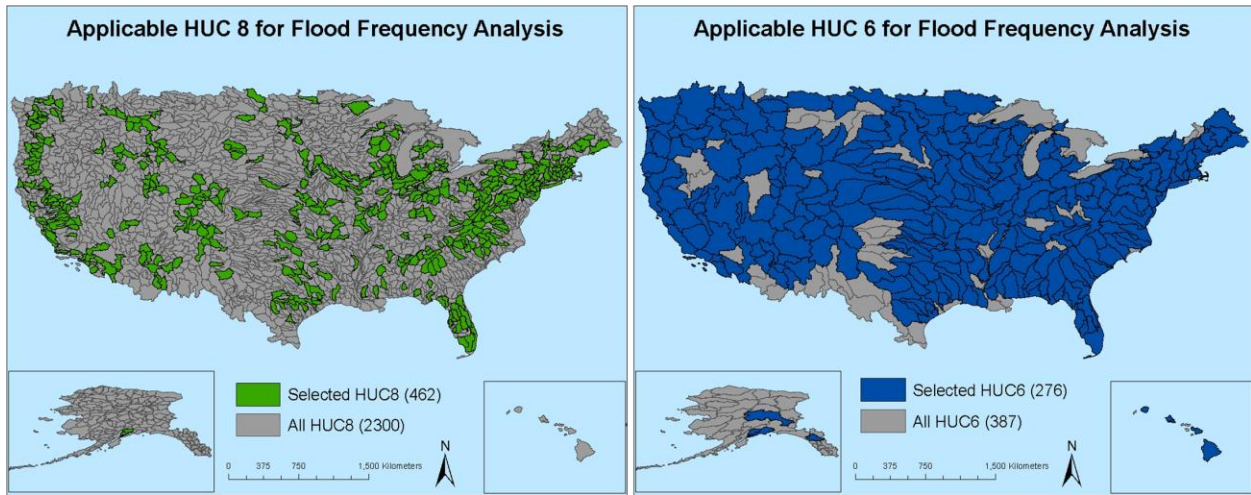
564

565

566

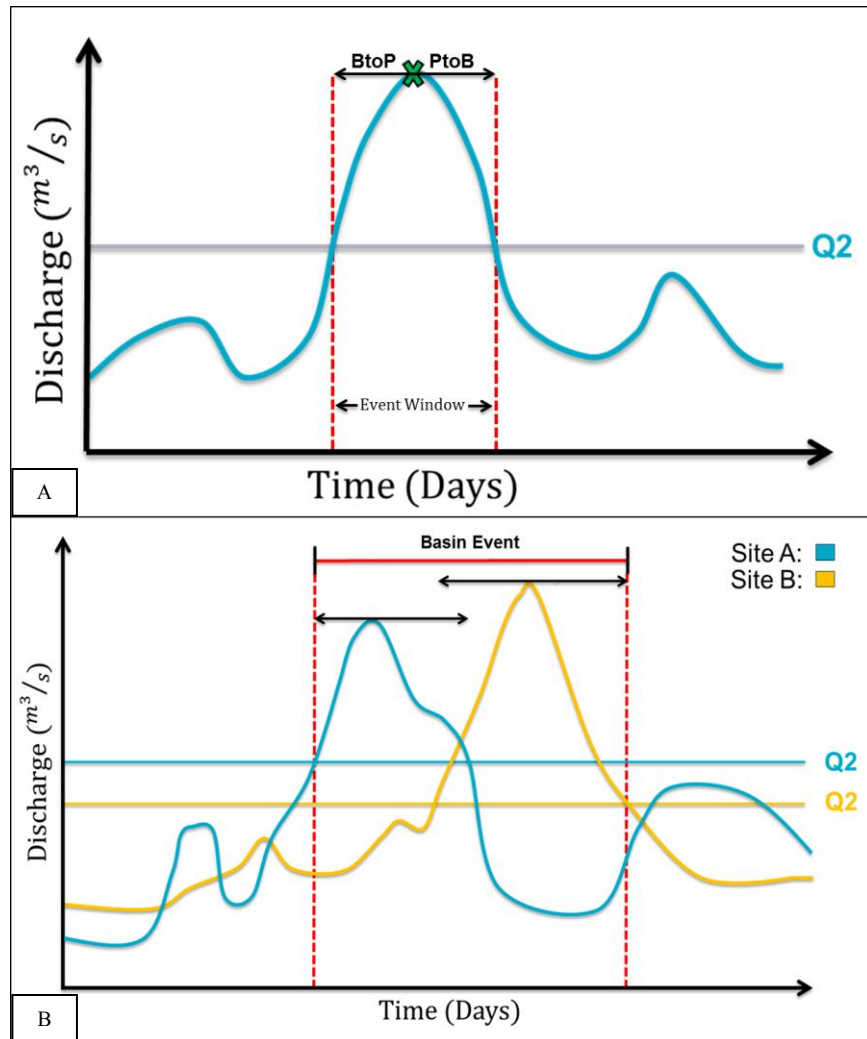
567

568 **Figures**



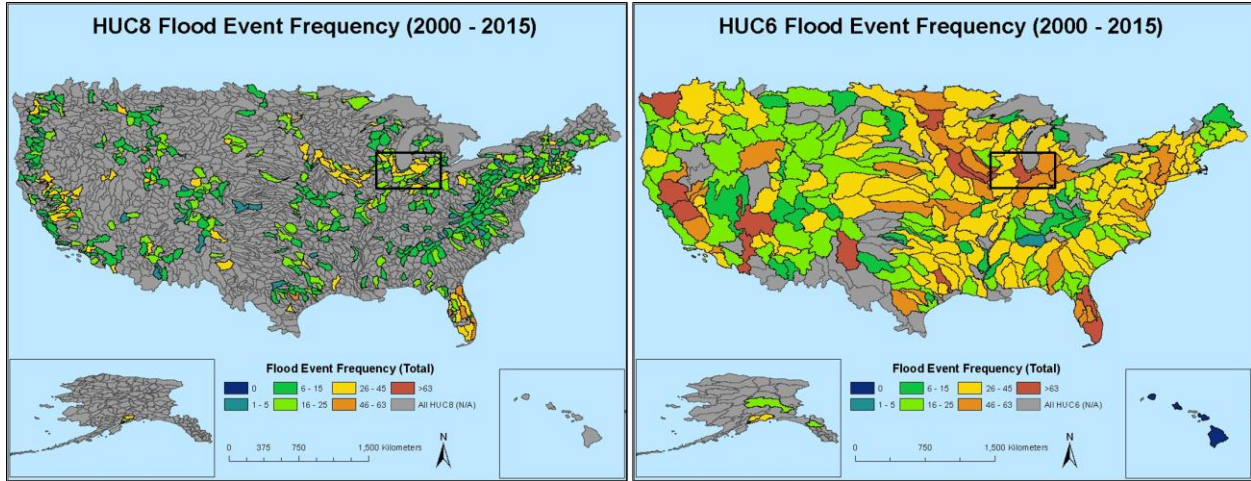
569 *Fig.1. A map of the selected HUC8 and HUC6*

570  
571



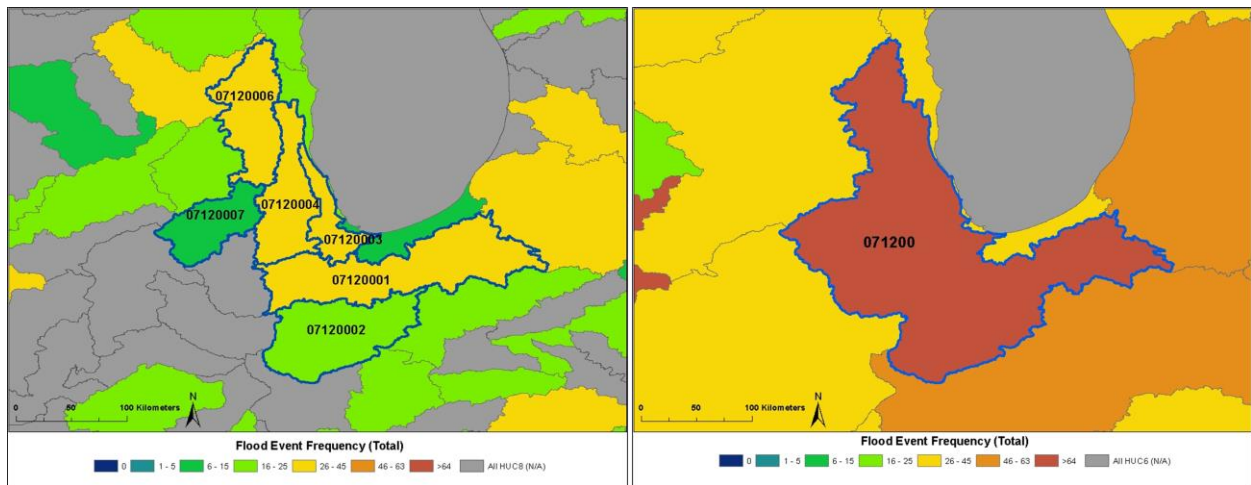
572 *Fig.2a. Site Event Identification (Top Panel) and Fig.2b. Basin Event Identification (Bottom Panel)*

573  
574



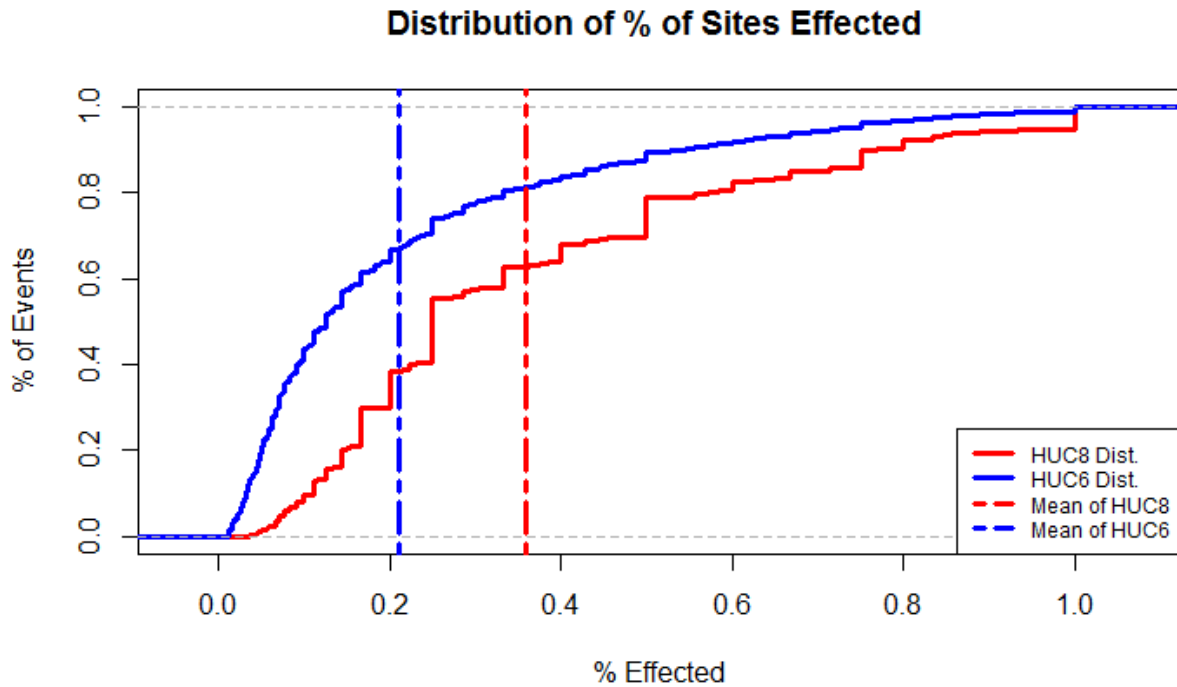
**Fig.3. HUC 8 and HUC6 Frequency Comparison, National**

575  
576  
577  
578



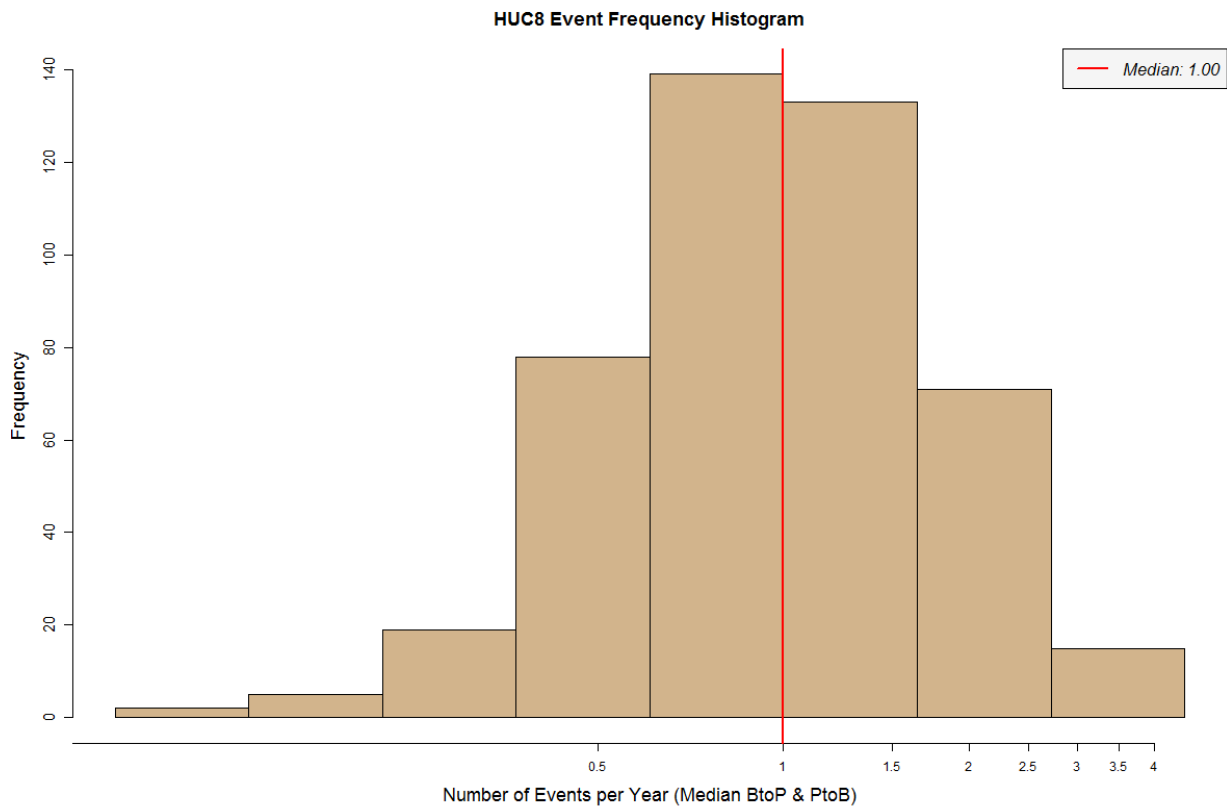
**Fig.4. HUC 8 and HUC 6 Frequency Comparison, Upper Midwest  
Blue Outline (HUC6: 071200, HUC8: 07120001---07120007)**

579  
580  
581  
582



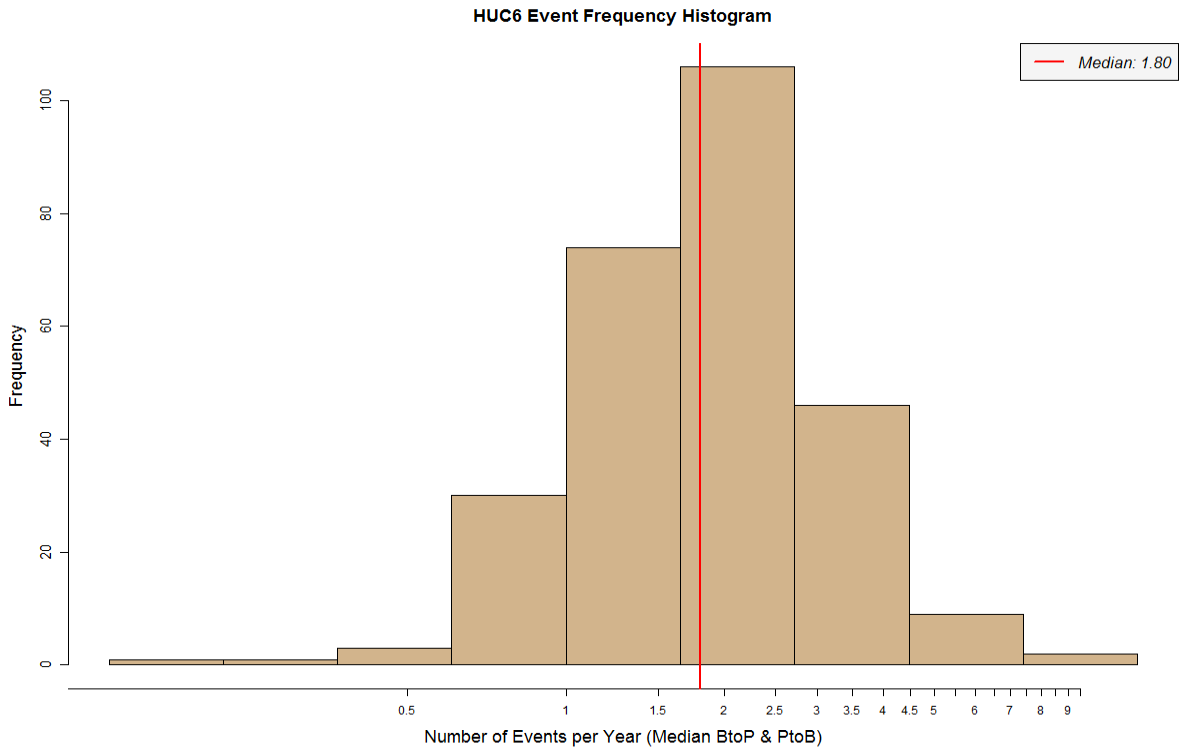
583  
584  
585  
586

*Fig.5. CDF of the percentage of sites impacted during each event within our catalog. Mean % of the entire distribution is noted and split by HUC.*



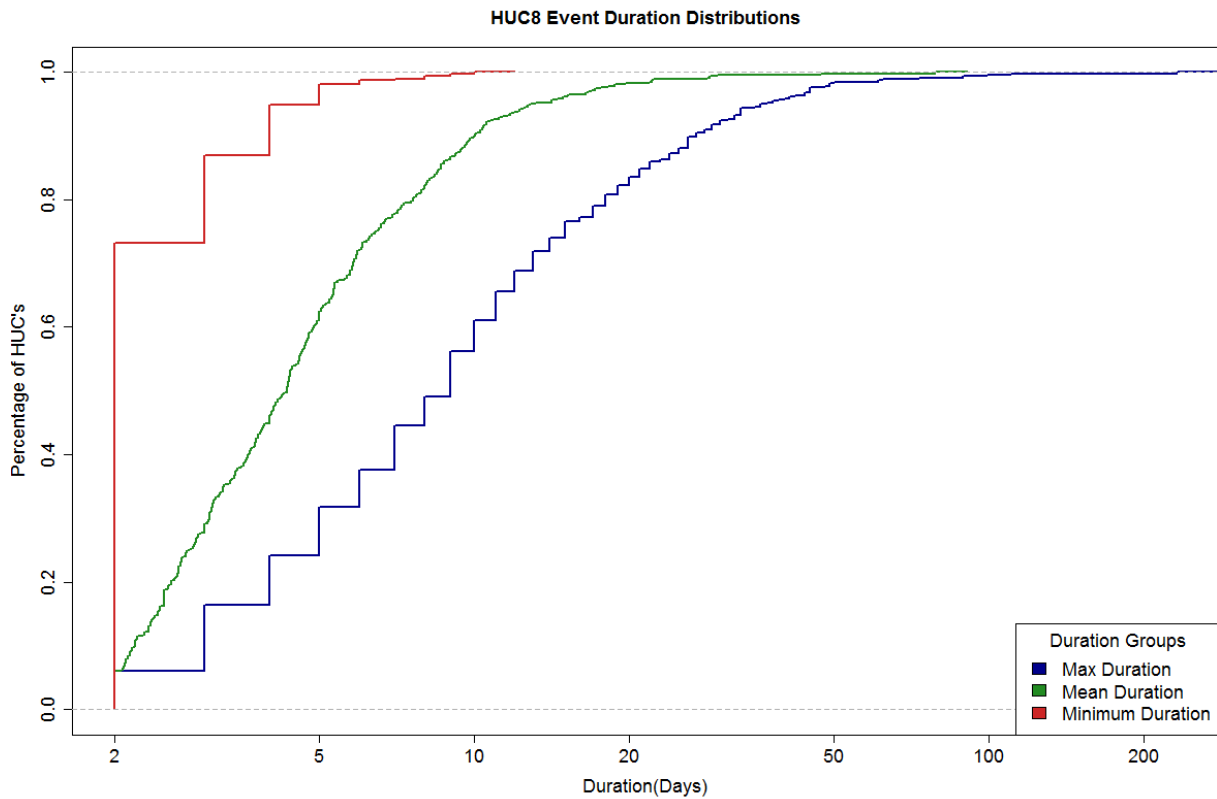
587  
588

*Fig.6. HUC8 Frequency Distribution*



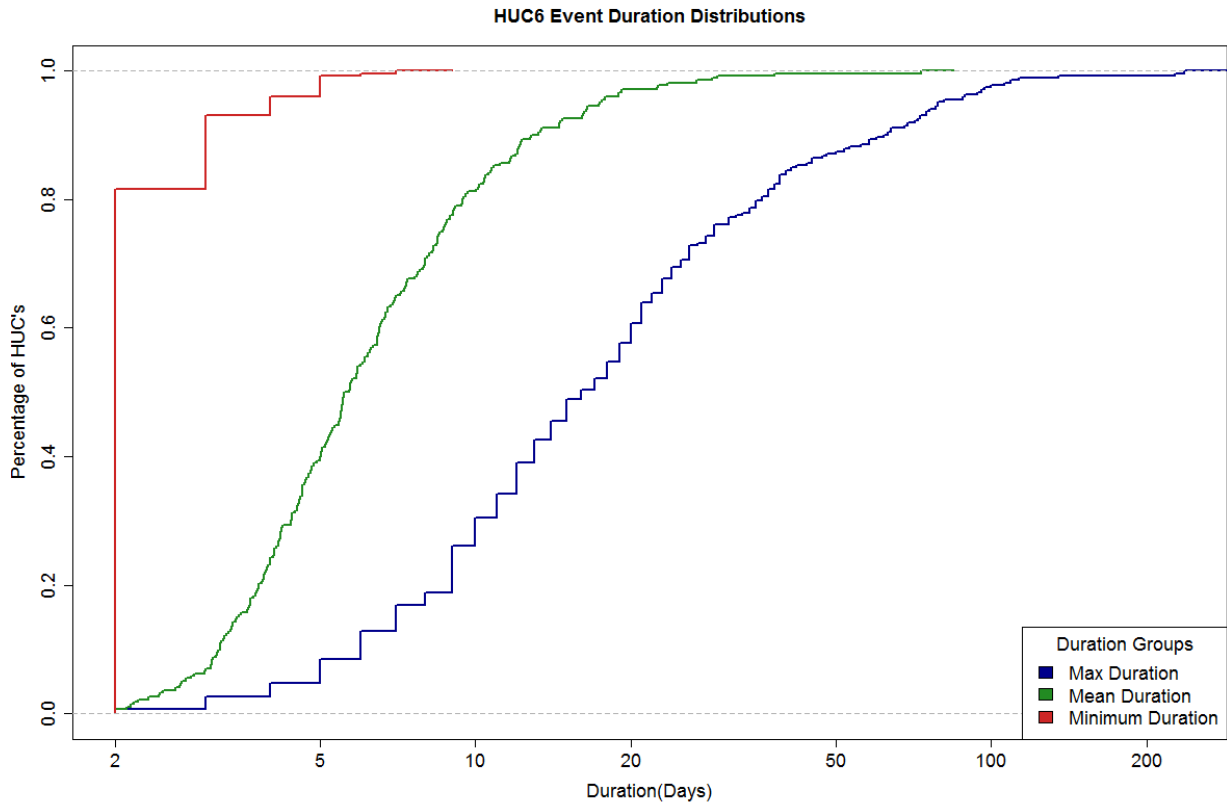
590  
591  
592

**Fig.7. HUC6 Frequency Distribution**



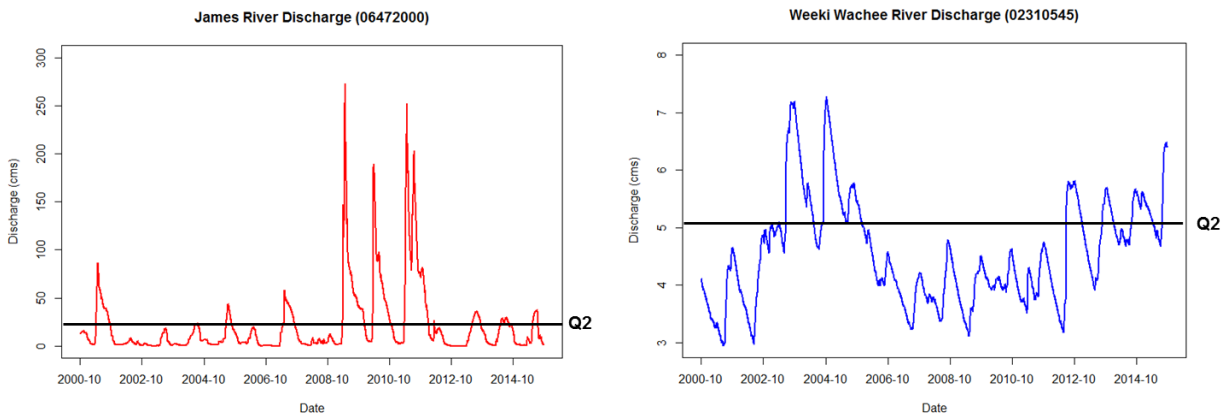
593  
594

**Fig.8. HUC8 Event Duration CDF**



595  
596  
597

*Fig.9. HUC6 Event Duration CDF*



598  
599  
600

*Fig.10. Example Sites for Event Duration Concerns*



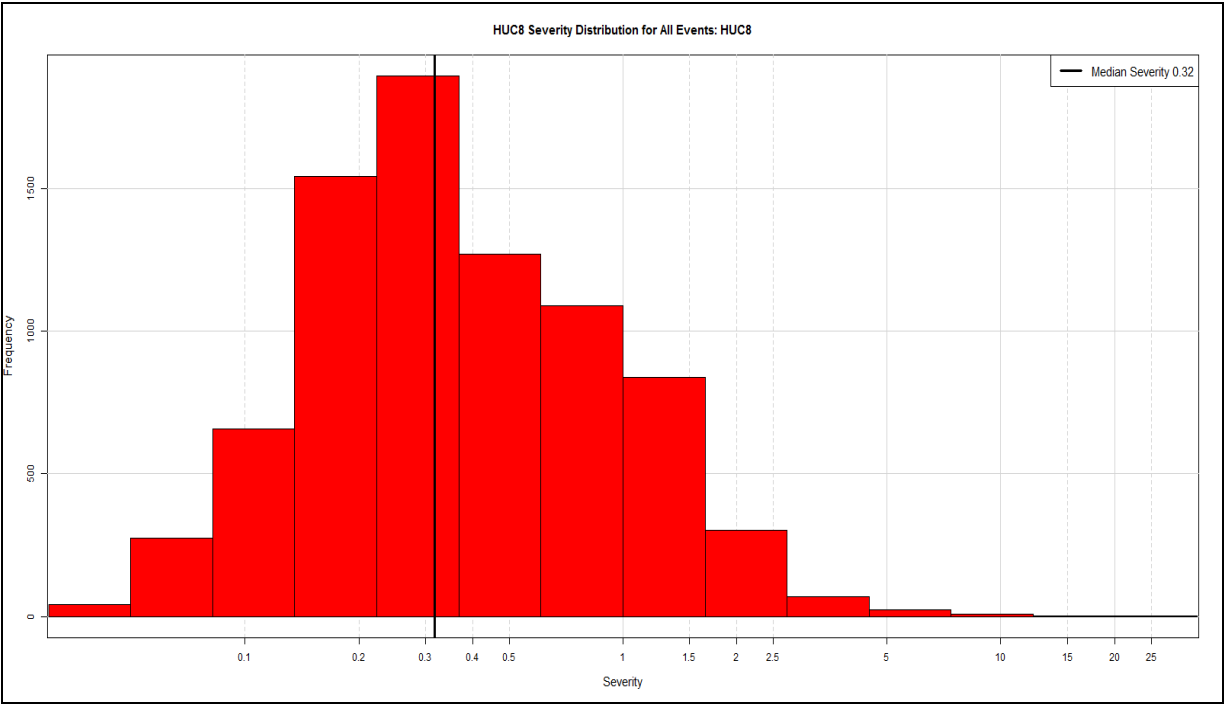


Fig.11. Severity Score Distribution

601  
602  
603

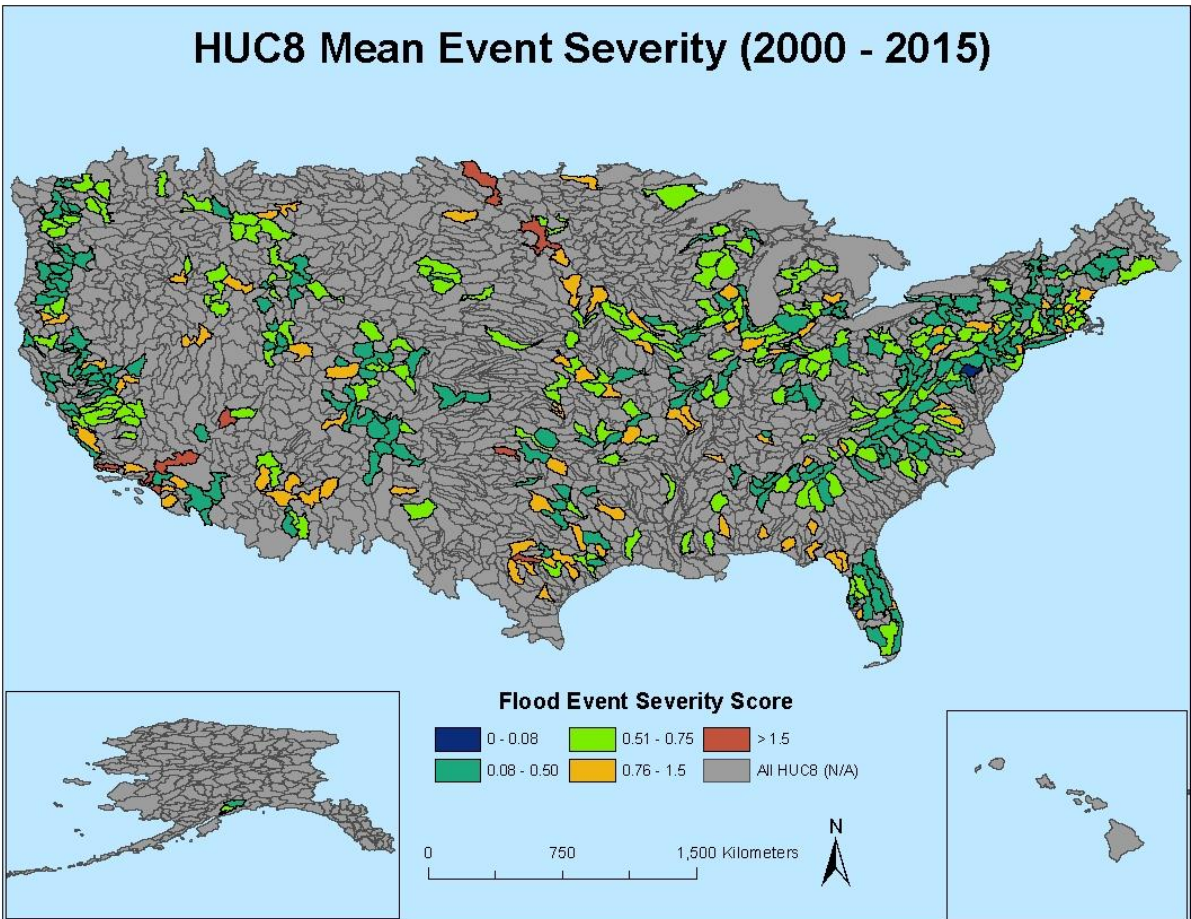
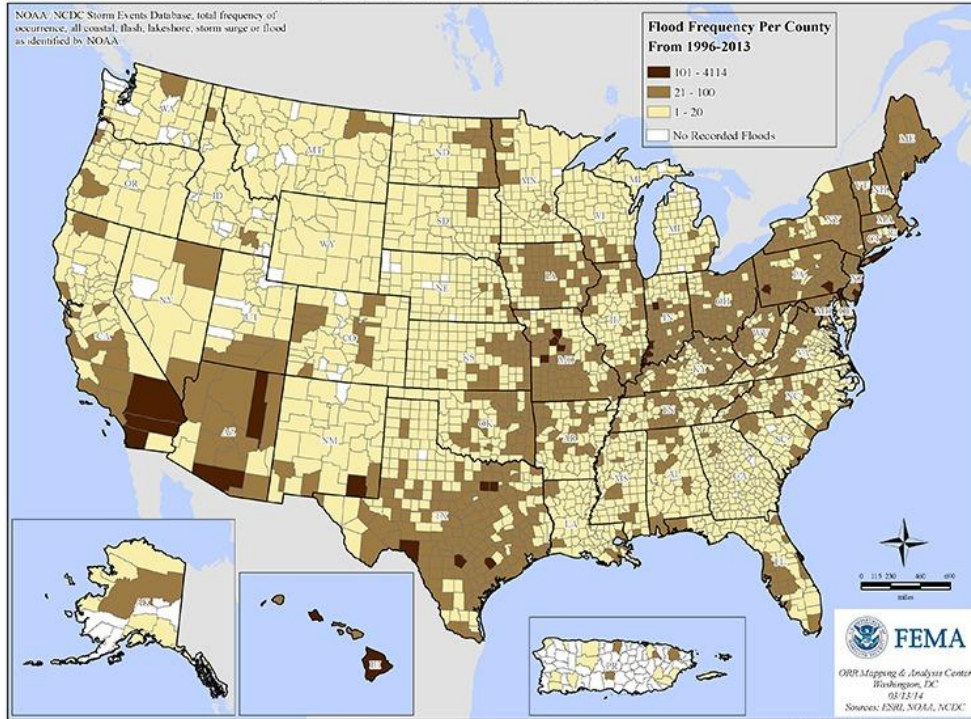


Fig.12. Regional Distribution of Severity

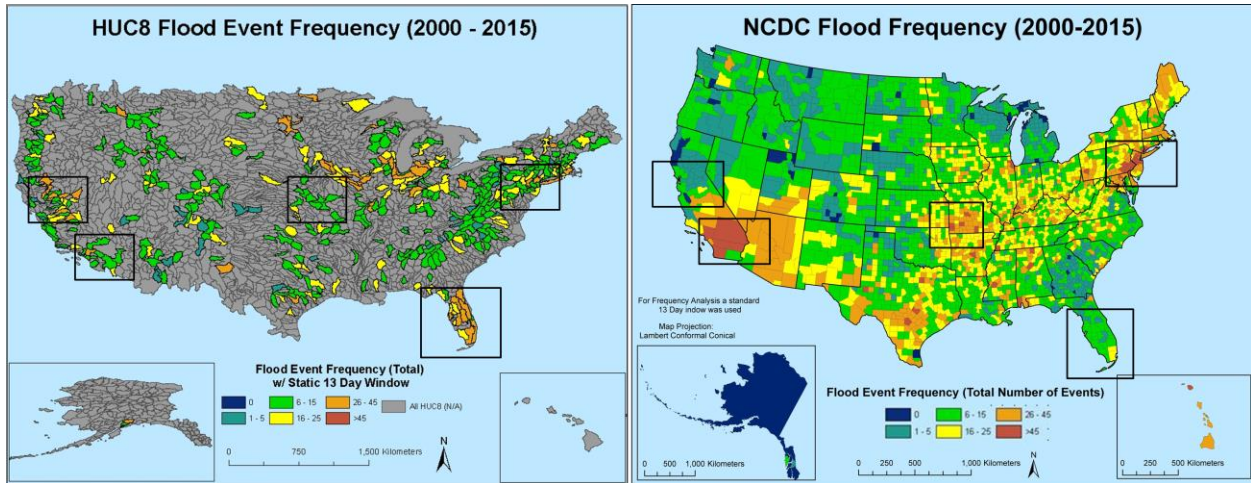
604  
605

**Frequency of Flood Events by County: 1996-2013**



**Fig.13 FEMA Flood Frequency Estimates**

606  
607  
608



**Fig.14. Frequency Comparisons with a 13 Day Window (NCDC & Daily Discharge)**

609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621

622 **References:**

- 623  
624 Bacova-Mitkova V, Onderka M. 2010. "ANALYSIS OF EXTREME HYDROLOGICAL  
625 EVENTS ON THE DANUBE USING THE PEAK OVER THRESHOLD METHOD." *Journal*  
626 *of Hydrology and Hydromechanics*, 58: 88-101. DOI: 10.2478/v10098-010-0009-x.  
627
- 628 Barredo, J. I., Saurí, D., and Llasat, M. C.: "Assessing trends in insured losses from floods in  
629 Spain 1971–2008", *Nat. Hazards Earth Syst. Sci.*, 12, 1723-1729, doi:10.5194/nhess-12-1723-  
630 2012, 2012.  
631
- 632 Black A. R. and Werritty A. "Seasonality of flooding: a case study of North Britain." *Journal of*  
633 *Hydrology*, Volume 195, Issue 1, 1997, Pages 1-25, ISSN 0022-1694,  
634 [http://dx.doi.org/10.1016/S0022-1694\(96\)03264-7](http://dx.doi.org/10.1016/S0022-1694(96)03264-7).
- 635 Dobour, J. C. and J. Noel. "A CLIMATOLOGICAL ASSESSMENT OF FLOOD EVENTS IN  
636 GEORGIA." *National Weather Service* (2005): n. pag. Web.
- 637 Doocy S., Daniels A., Murray S., Kirsch TD.: "The Human Impact of Floods: a Historical  
638 Review of Events 1980-2009 and Systematic Literature Review." *PLOS Currents Disasters*.  
639 2013 Apr 16, Edition 1. doi: 10.1371/currents.dis.f4deb457904936b07c09daa98ee8171a.
- 640 Gaffin, D. M. and D. G. Hotz. "PRECIPITATION AND FLOOD CLIMATOLOGY WITH  
641 SYNOPTIC FEATURES OF HEAVY RAINFALL ACROSS THE SOUTHERN  
642 APPALACHIAN MOUNTAINS." *NOAA/National Weather Service* 24.3 (2000): n. pag. Web.
- 643 Himmelsbach, I., Glaser, R., Schoenbein, J., Riemann, D., and Martin, B.: "Reconstruction of  
644 flood events based on documentary data and transnational flood risk analysis of the Upper Rhine  
645 and its French and German tributaries since AD 1480." *Hydrol. Earth Syst. Sci.*, 19, 4149-4164,  
646 doi:10.5194/hess-19-4149-2015, 2015.
- 647 Joyce, Christopher. "Federal Flood Insurance Program Drowning In Debt. Who Will Pay?" *NPR*.  
648 *NPR*, 1 Jan. 2014. Web. 05 Aug. 2016.
- 649 Kahana, R., Ziv, B., Enzel, Y., and Dayan, U.: "Synoptic climatology of major floods in the  
650 Negev Desert, Israel." *Int. J. Climatol.*, 22, 867–882, 2002.
- 651 Karnbaum K. and Wolfgang Kron, Dr. "What Is a Flood? Defining Flood Loss Occurrences for  
652 Reinsurance Purposes." *Munich Re* (2005): n. pag. Web.
- 653 Mallakpour, I., and G. Villarini. "The Changing Nature of Flooding across the Central United  
654 States." *Nature Climate Change* 5.3 (2015): 250-54. Web.  
655
- 656 Mallakpour, I., and G. Villarini, Investigating the relationship between the frequency of flooding  
657 over the central United States and large-scale climate, *Advances in Water Resources*, 92, 159-  
658 171, 2016.

659 "NATIONAL WEATHER SERVICE INSTRUCTION 10-1605." *National Climatic Data Center*  
660 (2007): n. pag. Web.

661 Nied M., Pardowitz T., Nissen K., Ulbrich U., Hundecha Y., Merz B., "On the relationship  
662 between hydro-meteorological patterns and flood types." *Journal of Hydrology*, Volume 519,  
663 Part D, 27 November 2014, Pages 3249-3262, ISSN 0022-1694,  
664 <http://dx.doi.org/10.1016/j.jhydrol.2014.09.089>.  
665

666 ORR Mapping and Analysis Center, Washington, D.C., March 13, 2014. Sources: ESRI, USGS.

667 "Overwhelming Risk: Rethinking Flood Insurance in a World of Rising Seas (2013)." *Union of*  
668 *Concerned Scientists*. N.p., Aug. 2013. Web. 05 Nov. 2016.

669 "The National Flood Insurance Program." *FEMA.gov*. FEMA, 20 May 2015. Web. 10 Aug.  
670 2016.  
671

672 Uhlemann, S., Thielen, A. H., and Merz, B.: "A consistent set of trans-basin floods in Germany  
673 between 1952–2002." *Hydrol. Earth Syst. Sci.*, 14, 1277-1295, doi:10.5194/hess-14-1277-2010,  
674 2010.  
675