A New Data-Driven Approach to Measuring Hurricane Risk

Duygu Pamukçu
Virginia Tech
duygu@vt.edu

Christopher W. Zobel
Virginia Tech
czobel@vt.edu

Andrew Arnette
University of Wyoming
aarnette@uwyo.edu

ABSTRACT

Improving disaster operations requires understanding and managing risk. This paper proposes a new data-driven approach for measuring the risk associated with a natural hazard, in support of developing more effective approaches for managing disaster operations. The paper focuses, in particular, on the issue of defining the inherent severity of a hazard event, independent of its impacts on human society, and concentrates on hurricanes as a specific type of natural hazard. After proposing a preliminary severity measure in the context of a hurricane, the paper discusses the issues associated with collecting empirical data to support its implementation. The approach is then illustrated by comparing the relative risk associated with two different locations in the state of North Carolina subject to the impacts of Hurricane Florence in 2018.

Keywords

Risk, Severity, Natural hazard, Analytic methods

INTRODUCTION

Disasters are catastrophic events that often cause many causalities and require the mobilization of significant numbers of resources to protect against their impacts (de Boer, 1990). Due to the increasing population of the world, and the increasing vulnerability of that population, these impacts are growing and require urgent attention. Rodriguez et al. (2011) argue that the political, economic and social considerations of global society force organizations to find more efficient ways of managing these destructive events. Assessing the characteristics of the events, and their possible destructive effects, is particularly important in the context of humanitarian relief operations, since they can have a significant impact on the ability to provide relief to those who need it in the wake of a disaster.

Resource allocation is a challenging problem in disaster management. Organizations spend their time and effort to find efficient ways of pre-positioning and distributing resources in order to provide better service to affected communities. Access to timely and appropriate information can play an important role in clarifying the situation and making appropriate decisions. Organizations should thus be aware of potential data sources to support such decision making. With this in mind, this study discusses a data-driven approach for quantifying disaster risk in terms of community characteristics and the inherent severity of a hazard event. We have a motivation of assessing the county-based risk of the hazard event to help organizations with their decision making, especially in efficient asset placement, in disaster management.

BACKGROUND

The social science literature includes many studies on measuring relative risk based on the likelihood of a hazard and the characteristics of the impacted community (Aerts et al., 2013, Willis et al., 2006). Arnette and Zobel (2019) adopt such an approach for quantifying risk as part of their model to optimize the placement of assets to support opening disaster shelters. Specifically, they use a common measure of risk that is a function of the probability of the occurrence of a hazard, *H*, the extent to which the population is exposed to that hazard, *E*, and the social vulnerability of the population (independent of any particular hazard), *V*:

$$R = H * E * V \quad (1)$$

By then letting H_{ij} represent the likelihood that a hazard of type j will occur in location i, where E_{ij} represents the corresponding level of exposure of the population in that location and V_i represents the corresponding vulnerability, Arnette and Zobel (2019) are able to calculate the total combined risk associated with multiple potential hazards in location i as follows:

$$R_i = \left(\sum_{j=1}^{n}\left(H_{ij}E_{ij}\right)\right)V_i \quad (2)$$

What is missing from this measure, however, is an explicit representation of the *severity* of these potential hazard events, independent of their impacts on the given location. This is important, since one would expect a more severe event to have a larger destructive impact on a given location than would a less severe event. This implies that if one were able to capture the relative severity of the hazard itself, then the varying levels of risk that a location is subject to under different circumstances could be characterized more effectively. This current study is a first step towards being able to explicitly incorporate a measure of exogenous severity into such an analysis.

Although the severity of hazard events is widely studied in the literature, the treatment of the concept varies depending on the type of study and its objectives. For example, Rodriguez et al. (2011) propose a decision support methodology to assist NGOs with the assessment of the consequences of disasters and provide a comparative analysis of different statistical learning techniques. The authors use a scenario-based approach to assess the severity of disaster in terms of its impacts by considering such factors as casualties, injured and homeless people, infrastructure damage, etc. De Boer (1990), in turn, combines the causes and the effects of a disaster to create a Disaster Severity Scale which includes a variety of attributes such as the effect on the surrounding community, the cause, the duration of that cause, the radius of the disaster area, and the time required by rescue teams for initial treatment. Alternatively, the study of de Boer et al. (1989) defines a Medical Severity Index of Disaster to support decision making in medical care supply chain operations in disaster situations. This index quantifies the severity of the incident by considering the level of injury and the requirement for treatment, along with other factors like disaster type, population, and location. Zhang and Huang (2018) also assign a severity value to all types of disasters, based on the number of deaths that occur, noting that droughts, extreme temperatures and earthquakes have the highest severity values and storms have lower severity but higher uncertainty.

The common thread among each of these research efforts is the assessment of severity by using measures that reflect the actual impacts on a population of a disaster event that is caused by a natural hazard. In contrast to this, our effort instead considers the inherent severity of the hazard itself, rather than that of the disaster, which should be measurable even in the absence of human impact. This is important because if one ultimately wishes to compare different communities' ability to resist the effects of a disaster, then it can be important to know the actual severity of the hazard itself, independent of its complex interaction with the communities. This would allow for more effectively comparing the relative contributions of different characteristics of the communities and their populations towards their ability to resist against and recover from a disaster.

MEASURING SEVERITY

In particular, our focus in this paper is on developing the capability to characterize the exogenous severity of a particular type of hazard event, a hurricane, in specific locations. Although the well-known Saffir-Simpson scale is frequently used to characterize the severity of an entire hurricane, based on the storm's maximum wind speed (NHC, 2018), different locations that are impacted by that same hurricane may experience very different conditions and thus not be subject to the same level of severity. Chouinard et al. (1997) discuss this general tendency to describe the severity of a hurricane as a function of the wind speed only, despite the presence of other potential measures of severity such as pressure deficit and maximum significant wave height. Aerts et al. (2013) also argue that not just high winds but also heavy rainfall and flooding due to storm surge are the main source of the destructive impacts of hurricanes. In addition, Schembri (2018) explicitly states that the main issue with Hurricane Florence, in particular, was rainfall rather than wind speed. Given these assessments, and the motivation for developing a localized measure of severity that can be related to risk in a specific location, as in equation (2) above, we focus on localized measurements such as rainfall, wind speed, wind direction, and flood warnings as potential indicators of hurricane severity.

This use of multiple indicators to represent the different aspects of the severity of a hurricane can easily be extended to the severity of other types of natural hazard such as that of an earthquake, which can be expressed in terms of magnitude, intensity, energy and acceleration (Gutenberg and Richter, 1942). In the long run, our hope is to define a generalizable approach for characterizing the core elements of severity that can be applied across different types of hazard. In this initial study, however, our focus is on hurricanes and on measuring their particular severity.

As a simple approach for measuring the localized severity, S, of a hurricane, we initially use a convex combination of a set of multiple indicator variables, *s*, each of which is normalized to maintain a consistent contribution:

$$S_i = \sum_{k=1}^{m} w_{ik} s_{ik} \tag{3}$$

In this case, *i* represents the specific location for which the hurricane's severity is being measured, *m* represents the total number of indicator variables chosen to be included in the measure, and *w* represents the weighted contribution of each indicator variable, as specified by the decision maker. Including weights in (3) allows for a decision maker to assign a greater relative importance to one or more of the indicator variables, and thus to prioritize the effects of wind over the effects of flooding, for example, under certain circumstances. Absent the motivation to make such a prioritization, however, we would expect each indicator variable to be weighted the same. Creating a single aggregated value for severity allows us to incorporate this severity measure into the measure of risk given in equation (2). In general, however, the set of indicators could be combined in some other way, or even conceptualized as a vector of different severity dimensions. In the context of this initial research effort, the straightforward formulation given in (3) simply allows us to establish how such a localized measure can be used to represent the complexity of the varying degrees of severity associated with a particular natural hazard.

EMPIRICAL DATA COLLECTION

We first explore the ability of a severity measure to characterize the local behavior of a hazard by applying our formulation of this measure to the analysis of a specific instance of a hurricane: Hurricane Florence, which struck the southeastern United States in September 2018. Florence was classified as a tropical storm on September 7th before strengthening to a major hurricane by September 10th and then weakening again to a Category 1 storm by the 13th of September. It made landfall near Wrightsville Beach, North Carolina on Friday, September 14th and then weakened further over the next few days as it moved inland and then northward. The following discussion details the process of collecting the empirical data upon which the analysis is based.

Weather data collection – We focused our data collection efforts for Hurricane Florence on the state of North Carolina. Historical weather data for each county in the state was retrieved from the National Oceanic and Atmospheric Administration (NOAA), using the *countyweather* package in R and including all data from September 7th to September 21st, 2018. The *countyweather* package, which requires an API key from NOAA, pulls all available weather data (both daily and hourly) from multiple weather stations. The core variables that it retrieves are temperature, wind speed, wind direction, and precipitation. Because the weather data is sparse for many counties during the specified date range, however, we ended up using only the daily average wind speed (in m/s) and the average precipitation (in mm).

In order to perform our preliminary analysis, we selected four specific counties with nearly complete data for the two variables of interest: these consisted of two adjacent counties from the coastal region (Carteret and Onslow) and two additional adjacent counties from the mountains (Buncombe and Henderson). Any missing data points in these counties were either estimated from hourly data, when available, or interpolated from the corresponding daily data (from the current day and the day before) in the adjacent county.

Figure 1. North Carolina State Map.
Source: https://commons.wikimedia.org/wiki/File:North_Carolina_counties.gif

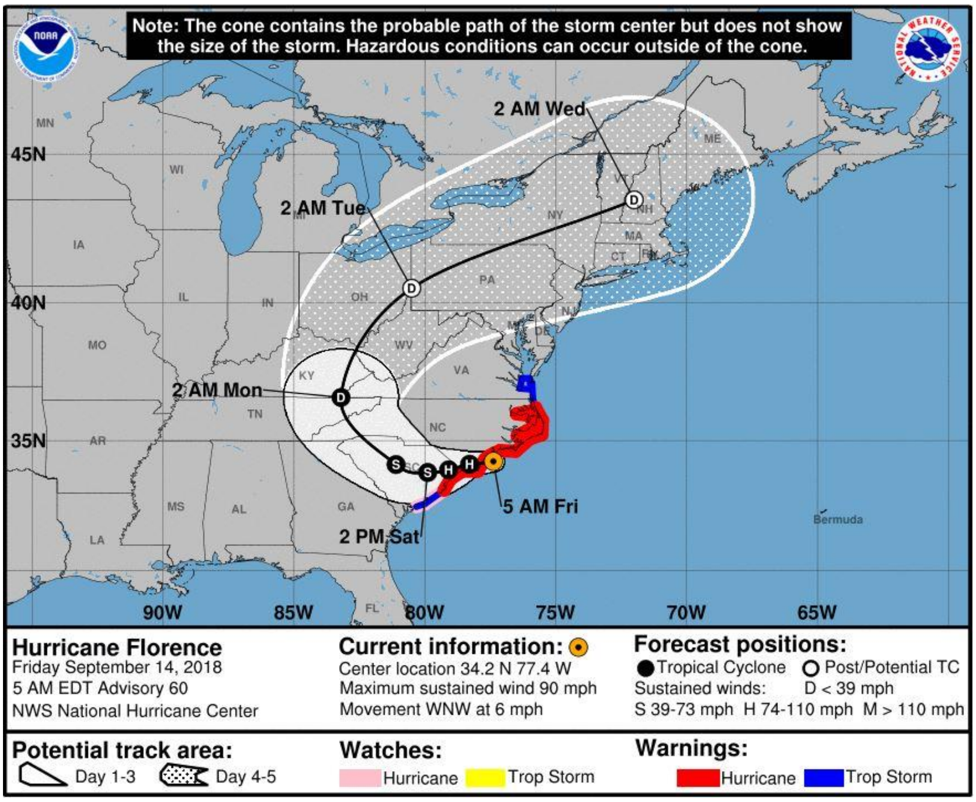

Figure 2. The National Hurricane Center map showing the forecast track of Hurricane Florence. Source: https://www.nhc.noaa.gov/archive/2018/FLORENCE_graphics.php?product=5day_cone_with_line

The motivation for selecting two different county groups was to allow for comparing the difference in severity between the counties that were first struck by Hurricane Florence and the counties over which it passed as it exited the state several days later. Figure 1 highlights the locations of the selected counties in red, and Figure 2 shows the path of the hurricane beginning with landfall and continuing until it finally exits the western part of the state 2-3 days later.

Vulnerability data collection – The Social Vulnerability Index (SoVI), which is a composite measure made up of more than 30 variables that capture important socioeconomic and built environment characteristics, is frequently used to quantify the social vulnerability of U.S. counties to hazards, based on their socioeconomic and demographic features (Cutter et al., 2003). In order to further incorporate the severity measure into the calculation of risk, this study used the most recent SoVI data set, which is based on 2010-2014 Census data.

Exposure data collection – For the exposure calculation within the risk function, we used the estimated base county population data for 2018 and multiplied it by a sheltering needs factor to estimate the proportion of the population needing assistance. Given that Mileti et al. (1992) suggest that such a sheltering needs factor can vary from a minimum of 0.05 to a maximum of 0.20, we set the factor at 0.20 for the coastal counties and at 0.05 for the counties in the western mountains, in order to reflect the inherent differences in the population's hurricane response between these two regions. To perform a more detailed analysis of actual sheltering behavior, rather than to simply illustrate the underlying modeling approach, both a displacement factor and a more accurate measure of sheltering need should be provided, as in Arnette and Zobel's (2019) work in Colorado and Wyoming. Unfortunately, such data is not yet readily available for the state of North Carolina at this time.

Table 1. County-level data

County	SoVI	Population
Onslow	-3.67	177,772
Carteret	1.03	66,469
Henderson	1.7	106,740
Buncombe	0.45	238,318

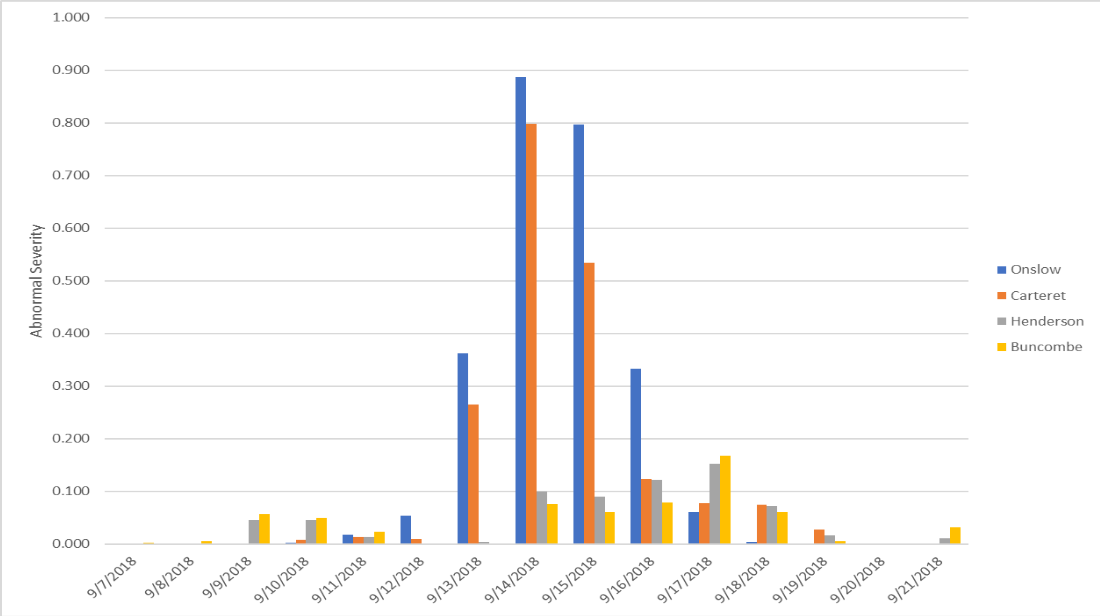

Figure 3. Abnormal severity related to Hurricane Florence

PRELIMINARY RESULTS

In order to generate a measure of the severity of the storm for each county during the specified time period, we calculated the average (normalized) wind speed and the average (normalized) precipitation over the five years previous to Hurricane Florence and created a weighted average of the two sub-measures, as in equation (3), using equal weights. This historical weighted average was used as a baseline against which the corresponding (normalized) values from September 2018 were compared. The amount by which the 2018 values exceeded the historical baseline was then taken as the actual measure of severity, with any values that fell at or below the historical weighted average being set to zero. Figure 3 presents the results.

It can clearly be seen in Figure 3 that September 14th is the day on which the hurricane reached its greatest severity along the coast, with slightly more severe weather being experienced in Onslow County. The severity in Carteret County, which is further north, drops off more quickly than that in Onslow County as the hurricane moves west and south. It can also be seen that the much reduced severity in Henderson and Buncombe counties grows in a similar way, but to a lesser extent and with a time lag of several days, as the depleted storm moves over that region. This behavior echoes expectations, given the forecast track of the hurricane as presented in Figure 2.

We then took the individual county-level severity values and incorporated them into an adjusted risk measure based on equation (2). Because we are considering only a single hazard event, and are characterizing the actual behavior of that event as it occurs, we assume a hazard probability of 1.0 and introduce the severity measure as the contribution of the hurricane to the local risk in each county:

$$R_i = S_i E_i V_i \tag{4}$$

The risk given by equation (4) thus reflects the relative extent to which each county was actually affected by the hurricane. Figure 4 subsequently illustrates the results of graphing the risk progression over the duration of the storm.

Because all three independent variables in equation (4) are normalized to the same [0, 1] scale, even a cursory comparison of Figures 3 and 4 clearly shows the significant influence of the severity measure on the measured risk value. The vulnerability also plays a significant role, however, as can be seen in the relative differences between the two coastal counties as well as those between the two mountain counties. Onslow County was subject to more severe conditions, yet its much lower population vulnerability leads to less overall risk during the height of the storm than Carteret County, despite its larger population. Similarly, despite its significantly lower vulnerability, Buncombe County's much larger population gives it a higher level of risk relative to that of Henderson County. In each case, it is obvious that the risk grows as the storm approaches and hits, and then drops as it moves away.

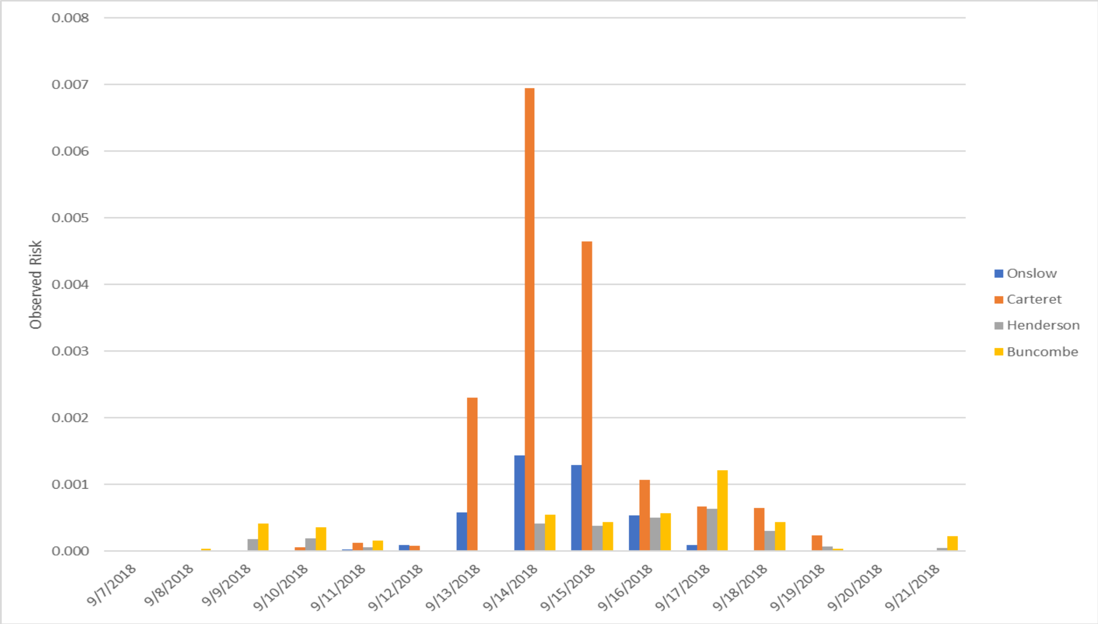

Figure 4. Observed risk related to Hurricane Florence

CONCLUSIONS

The purpose of this study is to provide a preliminary definition and illustration of a measure of the inherent severity of a natural hazard that can be incorporated into a localized measure of risk. Although a hazard can have a destructive impact on a given county, however, the inherent severity of that hazard is not the only factor affecting the risk to which that county's population is exposed. Indeed, the same level of severity may lead to different consequences in two different counties due to the relative vulnerability of the populations and the number of people who are actually exposed to the hazard event in a meaningful way.

It is important to recognize that this particular analysis was conducted in the context of a historical hazard event, and that the components of the severity measure were therefore given reasonably precise values. In order to calculate the risk for an upcoming event, not only would appropriate forecasts need to be available for the components' values, but also some notion of the probability of the event would need to be reintroduced into the discussion, since it may not affect a given community at all. This implies that the severity measure may also be effective in the context of slow-onset disasters, or at least disasters with less uncertainty, because it would be easier to make more accurate forecasts of future behavior. It also implies that scenario-based analyses on the severity of hazard events could give realistic insights about the possible outcomes of the hazard events that have possibility of occurrence in the future.

The focus of this paper was on a single natural hazard, and on the behavior of that particular hazard as it actually impacted several counties in North Carolina. This allowed us to assume a probability of occurrence of 1.0, and to effectively replace the hazard likelihood in equations (1) and (2) with the severity of the event. In general, however, if we wished to instead assess the future (forecast) severity and risk, we would need to also include the probability that the particular event would occur. This, in turn, would allow for considering the probability that different types of hazards might occur, and it would extend the applicability of the approach beyond just the set of indicator variables representing a single type of event.

Easily accessible weather data makes the severity calculation quick and easy to update, and the relative simplicity of both the severity (as defined above) and the risk function should make it very straightforward to extend the approach and adapt it to a variety of different problems. As a related future research direction, we are planning to include other accessible information, such as flood/storm warnings, into our severity measurement, and to consider other functional forms for the severity that may more effectively capture the relative contributions of its different characteristics in this context.

REFERENCES

Aerts, J. C., Lin, N., Botzen, W., Emanuel, K., & de Moel, H. (2013). Low probability flood risk modeling for New York City. *Risk Analysis*, 33(5), 772-788.

Arnette, A. N., & Zobel, C. W. (2019). A Risk-Based Approach to Improving Disaster Relief Asset Pre-Positioning. *Production and Operations Management*, 8(2), 457–478.

de Boer, J. (1990). Definition and classification of disasters: introduction of a disaster severity scale. *Journal of Emergency Medicine*, 8(5), 591-595.

de Boer, J., Brismar, B., Eldar, R., & Rutherford, W. H. (1989). The medical severity index of disasters. *Journal of Emergency Medicine*, 7(3), 269-273.

Chouinard, L. E., Liu, C., & Cooper, C. K. (1997). Model for severity of hurricanes in Gulf of Mexico. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 123(3), 120-129.

Cutter, S. L., Boruff, B. J. & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly*, 84, 242–261.

Gutenberg, B., & Richter, C. F. (1942). Earthquake magnitude, intensity, energy, and acceleration. *Bulletin of the Seismological Society of America*, 32(3), 163-191.

Mileti, D. S., Sorensen, J. H., & O'Brien, P. W. (1992). Toward an explanation of mass care shelter use in evacuations. *International Journal of Mass Emergencies and Disasters*, 10 (1), 25-42.

NHC (2018). Saffir-Simpson Hurricane Wind Scale. National Hurricane Center, NOAA. Retrieved from: https://www.nhc.noaa.gov/aboutsshws.php. Accessed on 2/13/2019.

NWC (2018). Historic Hurricane Florence, September 12-15, 2018. National Weather Service. Available at: https://www.weather.gov/mhx/Florence2018. Accessed on 2/14/2019.

Rodríguez, J. T., Vitoriano, B., Montero, J., & Kecman, V. (2011). A disaster-severity assessment DSS comparative analysis. *OR Spectrum*, 33(3), 451-479.

Schembri, F. (2018). Deadly storms break records, damage facilities. *Science*, 361(6408), 1172-1173. DOI: 10.1126/science.361.6408.1172.

Willis, H. H., Morral, A. R., Kelly, T. K., & Medby, J. J. (2006). Estimating terrorism risk. *Rand Corporation*, Santa Monica, CA.

Zhang, N., & Huang, H. (2018). Assessment of world disaster severity processed by Gaussian blur based on large historical data: casualties as an evaluating indicator. *Natural Hazards*, 92(1), 173-187.