



Using rapid damage observations from social media for Bayesian updating of hurricane vulnerability functions: A case study of Hurricane Dorian

Jens A. de Bruijn^{1,2}, James E. Daniell^{1,3}, Antonios Pomonis¹, Rashmin Gunasekera¹, Joshua Macabuag¹,
5 Marleen C. de Ruiter², Siem Jan Koopman^{4,5,6}, Nadia Bloemendaal², Hans de Moel², Jeroen C.J.H. Aerts^{2,7}

¹World Bank Group 1818 H St NW, Washington DC, 20433, USA

²Institute for Environmental Studies IVM, VU University, Amsterdam, The Netherlands

10 ³Geophysical Institute and Center for Disaster Management and Risk Reduction Technology, Karlsruhe Institute of Technology, Karlsruhe, Germany

⁴Department of Econometrics and Data Science, School of Business and Economics, Vrije Universiteit Amsterdam, The Netherlands

⁵CREATES, Aarhus University, Denmark

⁶Tinbergen Institute Amsterdam, The Netherlands

15 ⁷Deltares, Delft, The Netherlands

Correspondence to: Jens A. de Bruijn (jens.de.bruijn@vu.nl)

Abstract. Rapid impact assessments immediately after disasters are crucial to enable rapid and effective mobilization of resources for response and recovery efforts. These assessments are often performed by analysing the three components of risk: hazard, exposure and vulnerability. Vulnerability curves are often constructed using historic insurance data or expert judgments, reducing their applicability for the characteristics of the specific hazard and building stock. Therefore, this paper outlines an approach to the creation of event-specific vulnerability curves, using Bayesian statistics (i.e., the zero-one inflated beta distribution) to update a pre-existing vulnerability curve (i.e., the prior) with observed impact data derived from social media. The approach is applied in a case study of Hurricane Dorian, which hit the Bahamas in September 2019. We
20 analysed footage shot predominantly from unmanned aerial vehicles (UAVs) and other airborne vehicles posted on YouTube in the first 10 days after the disaster. Due to its Bayesian nature, the approach can be used regardless of the amount of data available as it balances the contribution of the prior and the observations.

1 Introduction

Natural hazards, such as tropical cyclones (Mendelsohn et al., 2012), floods (Winsemius et al., 2016) and earthquakes
30 (Bilham, 2009), affect millions of people and cost billions of dollars in damages every year. Their impacts are expected to increase further (Cutter et al., 2015) often attributed to factors such as climate change (Mora et al., 2018) and increasing exposure (Bouwer, 2011). Damage assessment immediately after a natural disaster is crucial to enable rescue and relief organizations to mobilize resources for response and recovery processes (Kryvasheyeu et al., 2016). However, detailed and accurate assessments often take months to complete (UNDP, World Bank, 2013). Therefore, the World Bank has recently



35 employed the Global RApid post-disaster Damage Estimation (GRADE) approach to rapidly estimate damages to physical
assets after major disasters (Gunasekera et al., 2018).

Damage estimations are commonly modelled using the three components of risk: hazard, exposure and vulnerability (Desai
et al., 2015). Hazard is defined as a potentially damaging event, exposure as the elements subject to damage and losses as a
40 result of a hazard and vulnerability as “the conditions determined by physical, social, economic and environmental factors or
processes which increase the susceptibility of an individual, a community, assets or systems to the impacts of hazards”
(UNISDR, 2016). Vulnerability and fragility functions are commonly used to model damage to buildings due to natural
disasters. These functions typically relate a measure of impact, such as wind speed (Pita et al., 2015), water depth (de Moel
et al., 2013) or ground motion (Li et al., 2013), to damage. These relations are mostly based on previous insurance claims,
45 experiments or expert judgment and are often only applicable to particular hazard characteristics and specific built
environments (Chung Yau et al., 2011; Douglas, 2007; Pita et al., 2013).

Observations, such as those from surveys (Wijayanti et al., 2017), can improve the accuracy of damage estimates (Douglas,
2007). However, affected areas are often difficult to access following a disaster (Bono and Gutiérrez, 2011), and human
50 resources are often limited (Koshimura et al., 2009). Data sources such as social media (Kryvasheyev et al., 2016),
unmanned aerial vehicles (UAVs) (Kim and Davidson, 2015) and other remote sensing techniques can provide detailed
observations of damage quickly during and after a disaster. However, the amount of data is heavily dependent on the
characteristics of the disaster area, such as the number of people that are able to use social media (de Bruijn et al., 2019; Yu
et al., 2018) and cultural differences (Cho et al., 2009).

55 A scientific challenge is to seek for methods that use observations from the affected area to improve vulnerability curves. An
example of such method is Bayesian analysis, which enables the updating of prior beliefs (e.g., beliefs based on expert
judgment) with observational data, irrespective of the amount of data available (Koutsourelakis, 2010). The balance between
prior information and observational data depends on the number of observations and the level of uncertainty in both the
60 observations and the prior beliefs.

Bayesian updating of fragility functions (i.e., the probability of exceeding a certain damage state) has been employed in
numerous studies. For example, Li et al. (2013) combined results from numerical simulations and experimental testing of
bridge substructures with Bayesian updating to obtain improved earthquake fragility functions. Mishra et al. (2017) updated
65 and quantified the uncertainty of analytical hurricane fragility functions for wood-frame buildings with experimental data. In
another study, a Bayesian framework was designed to create fragility functions for earthquakes (Koutsourelakis, 2010).
However, from a risk management perspective, vulnerability functions (where damage is represented as a damage ratio) are
more desirable, as they measure the actual capacity of the built environment in an event (Rossetto et al., 2015).



70 In this paper, we aim to improve existing vulnerability functions employed in GRADE assessments by updating these
 functions with post-disaster observations using Bayesian zero-one inflated beta regression (Ospina and Ferrari, 2010). The
 method is applied to a case study of tropical cyclone wind damage in the Bahamas caused by Hurricane Dorian in September
 2019. The observations were obtained from online media reports (i.e., YouTube) in the days following the disaster. This data
 includes some ground observations from driving cars but is mostly composed of observations from UAVs and videos shot
 75 from other airborne vehicles. In the remainder of this paper, we describe the general methodology (Sect. 2) and its
 application to the Bahamas (Sect. 3). We also discuss the applicability of the proposed method to other disaster types and
 different data sources.

2 Methodology

Bayesian updating or inference refers to the process of updating existing knowledge for a set of n parameters Θ , defined as Θ
 80 $= (\theta_1, \dots, \theta_n)$ and expressed in a prior distribution, defined as $P(\Theta)$, with new information X to find the posterior distribution,
 defined as $P(\Theta|X)$. Mathematically, we can express Bayesian updating as

$$P(\Theta|X) \propto P(X|\Theta) * P(\Theta). \quad (1)$$

The posterior distribution $P(\Theta|X)$ is obtained by multiplying the prior distribution $P(\Theta)$ by the likelihood distribution
 $P(X|\Theta)$, which is the probability of observing X given parameter set Θ , and a normalizing constant. The normalizing constant
 85 $P(X)$ is omitted here since this is automatically determined during the process of Gibbs sampling (see below; Gilks et al.,
 1996; Plummer, 2003b).

Following the Bayesian framework, the prior distribution $P(\Theta)$ for vulnerability must be defined. The vulnerability of the
 building stock can be expressed as a curve that maps wind speed as the explanatory variable (v_1, \dots, v_n) to a damage ratio
 90 from 0 to 1 (inclusive) as the response variable (y_1, \dots, y_n). Fragility curves of individual components of buildings (e.g., roof
 sheathing and nails) are widely regarded as following a cumulative lognormal distribution (Ellingwood et al., 2004; Lee and
 Rosowsky, 2005; Li and Ellingwood, 2006). Assuming identical fragility curves for individual components and
 independence of failure, a vulnerability curve for a building follows that same cumulative lognormal distribution (Holmes,
 1996):

$$y = \Phi\left(\frac{\ln(v/\alpha)}{\beta}\right) \quad (2)$$

where $y \in (0,1)$ is the damage ratio, α the median capacity of the building stock, β the logarithmic standard deviation of that
 capacity, $\Phi(\cdot)$ the cumulative probability density function for the standard normal distribution and v the sustained wind
 speed. While α and β can be expressed deterministically, we prefer to regard both parameters as uncertain and we express
 them as the random variables θ_1 and θ_2 , respectively.



100

For proportional data, the beta distribution is often used as the basis for the likelihood function $P(X|\theta)$ (Gupta and Nadarajah, 2004) because it supports a wide range of shapes on the interval (0, 1). Its probability density function, re-parameterized in terms of mean μ and precision φ , is given by Ferrari and Cribari-Neto (2004):

$$f_{\text{beta}}(y; \mu, \varphi) = \frac{\Gamma(\varphi)}{\Gamma(\mu\varphi)\Gamma((1-\mu)\varphi)} y^{\mu\varphi-1} (1-y)^{(1-\mu)\varphi-1}, \quad y \in (0,1) \quad (3)$$

105 where $\Gamma(\cdot)$ is the gamma function.

However, observations of the damage ratio can be true values of zero (i.e., no damage) and one (i.e., complete destruction), which are not supported by the beta distribution. In fact, such observations are more frequent because the assumption that the individual building components of buildings fail independently does not always hold. For example, houses are often completely destroyed (i.e., damage ratio of 1) due to the collapse of an important fundament (Keote et al., 2015) or complete displacement of the entire building (Shultz et al., 2005). By contrast, houses may be completely undamaged due to their environment or protection measures, such as shielding by other standing buildings (Keote et al., 2015). Using the zero-one inflated beta distribution enables us to explicitly model these 0 and 1 observations through probabilities π_0 and π_1 , respectively.

115

Therefore, for proportional response variables $y_i \in [0,1]$, Ospina and Ferrari (2010) propose a zero-one inflated beta regression. Here, the response variable (y_1, \dots, y_n) is modeled as a mixture probability function of π_0 , the probability that $y = 0$; π_1 , the conditional probability $Pr(y = 1 | y \neq 0)$; and the beta distribution with expected value μ_y and precision φ for the values between 0 and 1 (0 – 1). Its probability density function is given by:

$$f_{\text{ZOIB}}(y; \pi_0, \pi_1, \mu, \varphi) = \begin{cases} \pi_0 & \text{if } y = 0 \\ (1 - \pi_0)\pi_1 & \text{if } y = 1 \\ (1 - \pi_0)(1 - \pi_1)f_{\text{beta}}(y; \mu_y, \varphi) & \text{if } y \in (0,1) \end{cases} \quad (4)$$

By employing this distribution with proper parameterization (Sect. 3.3), we can model a process where it is highly probable that the damage ratio is 0 for low wind speed and 1 for high wind speed. For a wind speed in between, it is likely that y is modelled on the continuous scale (0-1) through the beta distribution. We base the equation for μ_y on Eq. (2) presented above:

$$\mu_y = \Phi\left(\frac{\ln(v/\theta_1)}{\theta_2}\right) \quad (5)$$

For probabilities π_0 and π_1 we specify a linear relationship with v on the logit scale as follows:

$$\pi_0 = \text{logit}^{-1}(\theta_3 + \theta_4 v) \quad (6)$$

$$\pi_1 = \text{logit}^{-1}(\theta_5 + \theta_6 v) \quad (7)$$



In Eq. (5-7), the parameters $\theta_1, \dots, \theta_6$ are assumed to come from normal distributions with some mean and standard deviation.

Finally, the prior distribution (Eq. 4-7) and parameters $(\theta_1, \dots, \theta_6)$ can be updated with some observations with the wind speed as the explanatory variable (v_1, \dots, v_n) and the corresponding response variables (y_1, \dots, y_n) using Gibbs sampling, which is a Markov chain Monte Carlo (MCMC) algorithm, see Gilks *et al.*, (1996), for a general treatment. For the purpose of Gibbs sampling, the predictor variable v is normalized such that $\tilde{v} \in (0,1]$ by dividing v by the maximum observed wind speed (i.e., $\max(v)$). All other input variables are scaled accordingly. Then, samples of the posterior distribution are generated using the Just Another Gibbs Sampler program (JAGS; Plummer, 2003a). We first use 1,000 iterations to tune the samplers (i.e., adaptation), 1,000 iterations as a burn-in to find the place where the Markov chain is most representative of the sampled distribution, followed by 100,000 iterations in three chains with a thinning of 100.

To verify the convergence of the Markov chains, we can present different diagnostics which are reviewed in, amongst others, Cowles and Carlin (1996) and Brooks and Roberts (1998). In particular, we concentrate on diagnostics based on distributional and autocorrelation statistics.

3 Case study of Hurricane Dorian

In this section, the methodology for updating vulnerability curves described above is applied to a rapid damage estimation of Hurricane Dorian in Grand Bahama and the Abaco Islands, the northernmost main islands of the Bahamas. First, the hazard (Sect. 3.1) and exposure components (Sect. 3.2) are briefly described. Then, we discuss the collection of observations for the vulnerability component and the Bayesian updating process (Sect. 3.3). Finally, these three components of risk are combined to obtain a final damage estimate (Sect. 3.4). It should be noted that all data sources had to have been collectible in the first 10 days following the first landfall of the Hurricane in the Bahamas to be eligible for the rapid damage estimate.

3.1 Hazard

On 24 August 2019, Tropical Storm Dorian formed over the Atlantic Ocean and began tracking through the Windward Islands and the U.S. Virgin Islands towards the Bahamas (Avila *et al.*, 2020). On 1 September, Dorian reached Category-5 strength with maximum sustained wind speeds exceeding 300 km/h. At 16:40 UTC that day, Dorian made landfall at Elbow Cay, Great Abaco, in the Bahamas. During its passage over the Great Abaco and Grand Bahama islands, the weakening of the nearby high-pressure area caused the hurricane to lose its steering currents and therefore significantly slowing its forward speed to 2 km/h, at times even coming to a standstill. Dorian remained near-stationary for 36 hours, until the hurricane started moving north-northwestwards towards North Carolina (USA). Dorian dissipated on 8 September near Canada.



160 In this study, meteorological conditions during Hurricane Dorian’s passage over the Bahamas is taken from the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al., 2010). Using the position of the eye, the minimum pressure, maximum wind speed, and size of the eye (i.e., radius to maximum winds) the 2D wind field is constructed by applying the parametric approach of Holland (1980), which was further refined in by Lin and Chavas (2012). To obtain the wind field at 10 meter above surface level (Fig. 1) a reduction factor of 0.85 is applied (Powell et al., 2005).

165 3.2 Exposure

To determine the exposure of the residential buildings on the Abaco and Grand Bahama Islands, we consulted the 2000 and 2010 Population and Housing Census of the Bahamas (Department of Statistics, 2002, 2012). The census contains information about the housing stock. Specifically, it contains data on residential buildings and occupied and vacant dwelling units for each settlement and enumeration district in the Abaco Islands and each supervisory district in Grand Bahama, as well as the number of bedrooms and the total annual household income for each household size. In the Abaco Islands (and to a lesser extent in Grand Bahama), the proportion of vacant housing stock is relatively high as the islands are a tourist destination, and many homes are owned or rented by vacationers. The Abaco Islands also has many migrant communities (mainly working in the service sectors for the tourism industry), who reside in low-quality housing in several informal settlements, such as the Mudd and the Pigeon Peas settlements. Other settlements are occupied by the local Bahamian population, while tourists and foreign citizens, often reside in high-value homes and resorts.

To account for this heterogeneity all settlements and supervisory districts (“regions”) were mapped individually and the value of residential buildings within three building classes (low-, medium- and high-quality) was estimated for each region. We first estimated the number and area of buildings per building type in each region. To do so, we consulted building footprints from OpenStreetMap (and found 16,100 and 12,500 building footprints in Grand Bahama and the Abaco Islands respectively). We estimated the housing stock in 2019 by projecting data from the 2000 and 2010 census onto 2019, taking official population projections into account (Department of Statistics, 2015) and estimated the floor area of the dwellings using data on the number of bedrooms per dwelling. We also consulted the Bahamas 2013 Household Expenditure Survey (Department of Statistics, 2016), which provided information about household consumption quintiles and housing conditions, such as type of dwelling, number of rooms, bedrooms, period of construction of the dwelling, type of tenure, type of construction material used for the outer walls, the roof cover and the floors.

Next, we estimated the unit cost of construction of the building classes based on housing prices provided by real estate agents, building contractors and a census of informal settlements in the Abaco Islands (Shanty Town Task Force, 2018). Finally, we consulted the Annual Building Construction Statistics Reports (e.g., Department of Statistics, 2018), which contain data on the investment values of newly constructed buildings. Figure 2 shows the reported monetary value of low-quality residential buildings per region (Fig. A1 and A2 for values of medium- and high-quality buildings; note the adjusted



scales). Note that these estimations do not include building content. Since no official maps of the regions were available, we determined the coordinates of the population centres for each region. All maps are presented as Voronoi diagrams (i.e., partitioned into regions closest to each centre point) based on these centroids.

3.3 Vulnerability

To derive event-specific vulnerability curves, we aimed to update vulnerability curves derived from previous hurricane observations in similar built environments using damage observations for individual buildings in the affected area. Therefore, we analysed all 498 YouTube videos that were listed when we searched for “*Bahamas Dorian*” and that were posted in the 10 days after the first landfall of Hurricane Dorian in the Bahamas (September 1st–9th). We then analysed all videos that 1) showed an overview of an area or a row of buildings (to ensure the sample was as representative as possible), 2) that we were able to geotag (i.e., locate) and 3) that showed buildings that did not appear to have undergone extensive flood damage. This resulted in a set of 15 videos (Appendix B), from which we extracted 732 buildings. Figure 3 depicts two examples of buildings extracted from the videos. By comparing the footage with satellite imagery of the area before the hurricane, we ensured that all buildings in an observed area or row of buildings were analysed, including those that completely disappeared in the storm.

Then, the damage ratio [0-1] and building class (i.e., low-, medium, and high-quality) were estimated for each building. Based on experience in post-disaster damage assessment in insurance, economic damage ratios were estimated based on the damage seen in the value of subcomponents of the structure and their relative values and interactions as a whole, see Massarra et al. (2019). In some areas, especially those with low-quality houses, it was difficult to extract an image of each individual building due to the large amount of destruction and displacement. In such cases, we estimated the number of buildings per level of damage from pre-event satellite imagery.

Next, we derived a prior (i.e., a vulnerability curve based on pre-existing knowledge; Sect. 2) for each building class by estimating the parameters based on expert judgment of the strength of the buildings in similar built environments. Such curves have fitted the results of previous PDNAs within the Caribbean for stronger wind events for economic damage, and provide similar smoothed curves to existing models in the region such as those presented in the UNDRR’s Global Assessment Report on Disaster Risk Reduction (UNISDR, 2015) for developed locations. Figure 4 (left) shows the parameters α and β of these curves for the three housing qualities and Fig. 4 (right) shows the associated damage ratio for low-quality buildings in the affected region.

We then set the priors (Fig. 5) for μ_{θ_1} and μ_{θ_2} (Eq. 5) using the specified values of α and β for each building class; σ_{θ_1} was set to 15 using the uncertainty range expressed in the fragility curves for a specific damage state (i.e., the range of wind gust speed that causes a specific damage state) in HAZUS (a tool for analysing natural hazards in the United States) for a similar



built environment (Vickery et al., 2006) and σ_{θ_2} was set to 0.03 to allow for uncertainty of the building typology. The values of the means and standard deviations of $\theta_3, \dots, \theta_6$ (Eq. 6, 7) were set such that the probability π_0 is near one when μ_y is near zero and π_0 near zero when μ_y is near one. Conversely, π_1 is set such that its value is near one when μ_y is near one and π_1 near zero when μ_y is near zero. This is explained in full in Appendix C. For precision parameter φ , we use an uninformed
 230 uniform prior.

For each building's observed damage ratio $(y_1, \dots, y_n) \in [0,1]$, the maximum sustained wind speed (v_1, \dots, v_n) was obtained using the location of the building (Sect. 3.1). Finally, using these observations we performed Gibbs sampling to obtain the posterior distribution (Fig. 5; Sect. 2). In Fig. 5, the upper row (a-c) shows the prior vulnerability curves, whereas
 235 the lower row (d-f) shows the posterior curves. From left to right the columns represent the vulnerability curves for low- (a,d), medium- (b,e) and high-quality (c,f) buildings. The red, blue and green colours denote respectively π_0 , π_1 and μ_y . The solid line represents the median (P_{50}), the dashed lines (P_{25-75}) represent the 25th and 75th percentile and the shaded area represents the 10th-90th percentile range.

240 Comparing the top and the bottom row, the curves for all building types have shifted substantially showing the result of the Bayesian updating. Note that for the posterior vulnerability curves for low-quality buildings (bottom left; d) it appears that there is a significant probability that the building is entirely destroyed (green curve). However, this is not the case since π_1 is the conditional probability that the damage ratio is one given that the damage ratio is not zero ($Pr(y = 1 | y \neq 0)$; Eq. 4).

245 Figures D1, D2 and D3 display the values per iteration, density plot and autocorrelation for $\theta_1, \dots, \theta_6$ for each building quality class. While most sequences of generated parameters in the MCMC have low autocorrelations, some parameters do show high autocorrelations (i.e., θ_3 and θ_4 for low- and medium- quality buildings and θ_1 and θ_2). This is likely caused by the absence of data information for these parameters, together with a possibly large disagreement between the prior and observed data.

250

Figure 6 shows the posterior damage ratio per district for low-, medium- and high-quality buildings.

3.4 Damage estimation

Finally, the three components of risk were combined (i.e., hazard \times exposure \times vulnerability) to obtain a damage estimation for residential buildings (Table 1; Fig. 7) of 1056 million USD using the prior vulnerability curves versus an estimate of 658
 255 million USD using posterior curves (i.e., ~38% lower). It should be noted that we used the coordinates of the population centre determined for each region (Sect. 3.2) to extract the wind speed data (Sect. 3.1).



4 Discussion and conclusions

In this paper, we present a framework that uses Bayesian updating with social media data (YouTube videos) to create event-specific vulnerability curves. This framework uses the zero-one inflated beta distribution, which allows us to use post-disaster observations to create vulnerability curves that have been adjusted for local hazard and building characteristics. We demonstrate their application in a rapid damage assessment of structural damage to buildings caused by Hurricane Dorian in the Bahamas. In our estimation, wind damage to residential buildings is ~38% lower compared to that calculated using pre-existing vulnerability curves (i.e., the priors). The largest relative differences were found for medium- and high-quality buildings, which we argue are most likely designed to be designed according to strict building codes (Ministry of Works & Utilities, 2003).

However, using social media data to assess building damages has several limitations. Observations from online media are biased, and some demographic groups have easier access to internet resources than other groups (Duggan and Brenner, 2013). In addition, observations tend to focus on the most impacted areas, as these are more newsworthy (Miles and Morse, 2007). While we have aimed to reduce this bias by only including footage that showed relatively large areas, we found very little footage of the less severely impacted parts of the islands, such as the northern and southern tips of the Abaco Islands. Moreover, the large spread in the observations (Fig. 5) shows that vulnerability is a complex concept. The vulnerability of a single building or part of building to a specific hazard is determined by many factors. In this paper, we only considered sustained wind speed. However, rainfall patterns and other environmental factors could also be important (Hatzikyriakou et al., 2016; Knapp et al., 2010). In addition, while we used three different building classes, the real variation in the strength of these buildings cannot be captured by three relatively simple curves.

The variation in building classes was also difficult to capture. For our case study, we based our classification of damaged buildings on post-disaster imagery. While we aimed to deduce the building quality class from the building structure rather than from the damage to the building, it is likely that bias was introduced in the classification. A better approach would be to use pre-disaster imagery (e.g., Google Street View or Mapillary) or, even better, detailed construction data per building. Unfortunately, these were unavailable for the Grand Bahamas and the Abaco Islands.

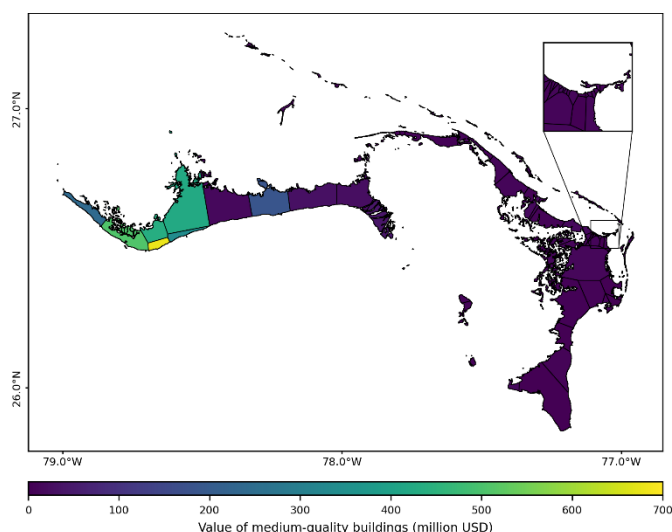
This was further complicated by the limited availability of vulnerability curves for the specific built environment. The estimated priors are based on a combination of data sources and expert judgement. This likely caused a relatively large disagreement between priors and observations for some parameters, resulting in a high autocorrelation (Fig. D1, D2 and D3). This calls for the establishment of a database of vulnerability curves for hurricane winds that considers a wide range of building typologies and strengths across a wide spectrum of wind intensities.



290 We applied the method using observations from online media. However, point observations from other sources could also be
 used. For example, survey data collected by experts or insurance claims would likely be more reliable. However, the
 availability of such observations from such sources during the first days following a disaster is generally limited, which
 reduces their applicability for rapid post-disaster damage assessments. Our method could be applied to other hazard types,
 such as floods and earthquakes. However, building damage caused by standing water may be more difficult to observe from
 295 pictures than structural damage caused by earthquakes. Therefore, for floods additional data, such as data from surveys or
 insurance reports, may be required.

5 Appendices

Appendix A



300 **Figure A1: Value of medium-quality buildings per region.**

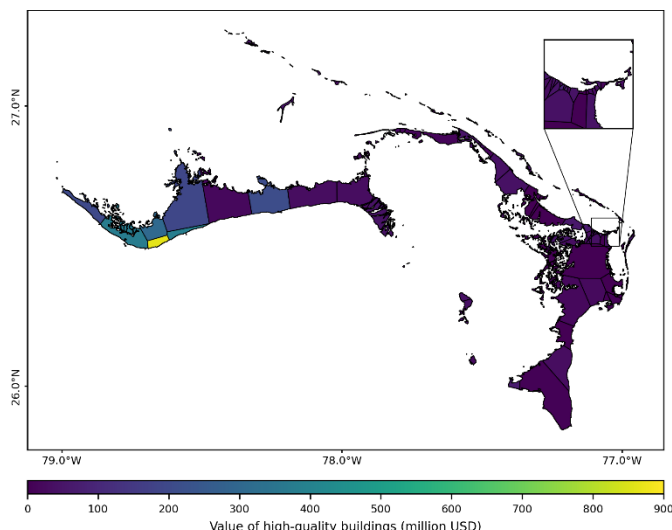


Figure A2: Value of high-quality buildings per region.

Appendix B

- <https://www.youtube.com/watch?v=rq95eJWxpd8>
- <https://www.youtube.com/watch?v=beXg9egFcAs>
- https://www.youtube.com/watch?v=U0omTvsIr_U
- <https://www.youtube.com/watch?v=SuzXtldZqvg>
- <https://www.youtube.com/watch?v=8QoqtB6HPMY>
- <https://www.youtube.com/watch?v=SN4jgJX0OP8>
- <https://www.youtube.com/watch?v=hvCQtLWW-y4>
- <https://www.youtube.com/watch?v=1l1ZzCpeILs>
- <https://www.youtube.com/watch?v=SrUfwnX-UjI>
- <https://www.youtube.com/watch?v=Ar2BdIEI59w>
- <https://www.youtube.com/watch?v=9PypGi8M29Q>
- <https://www.youtube.com/watch?v=mpgGw3iyBPU>
- <https://www.youtube.com/watch?v=AdHombKmG78>
- https://www.youtube.com/watch?v=_tFfnIGq2qE
- <https://www.youtube.com/watch?v=E-zdt-fYLIg>

Appendix C

μ_{θ_3} and μ_{θ_4} are chosen such that π_0 (i.e., the probability that the damage ratio is zero) is 0.99 where μ_y is 0.01 and π_0 is 0.01 where μ_y is 0.05. In simpler terms this means that the probability that the response variable y_i is a true zero, is near-one when μ_y is also near-zero, and near-zero otherwise (P_{50} or median in Fig. 5). The standard deviation is set at 10% of the mean. Likewise, μ_{θ_5} and μ_{θ_6} are chosen such that π_1 (i.e., the probability that the damage ratio is one) is 0.01 where μ_y is 0.95 and π_1 is 0.99 where μ_y is 0.99. Likewise, the standard deviation is set at 10% of the mean.



325 Appendix D

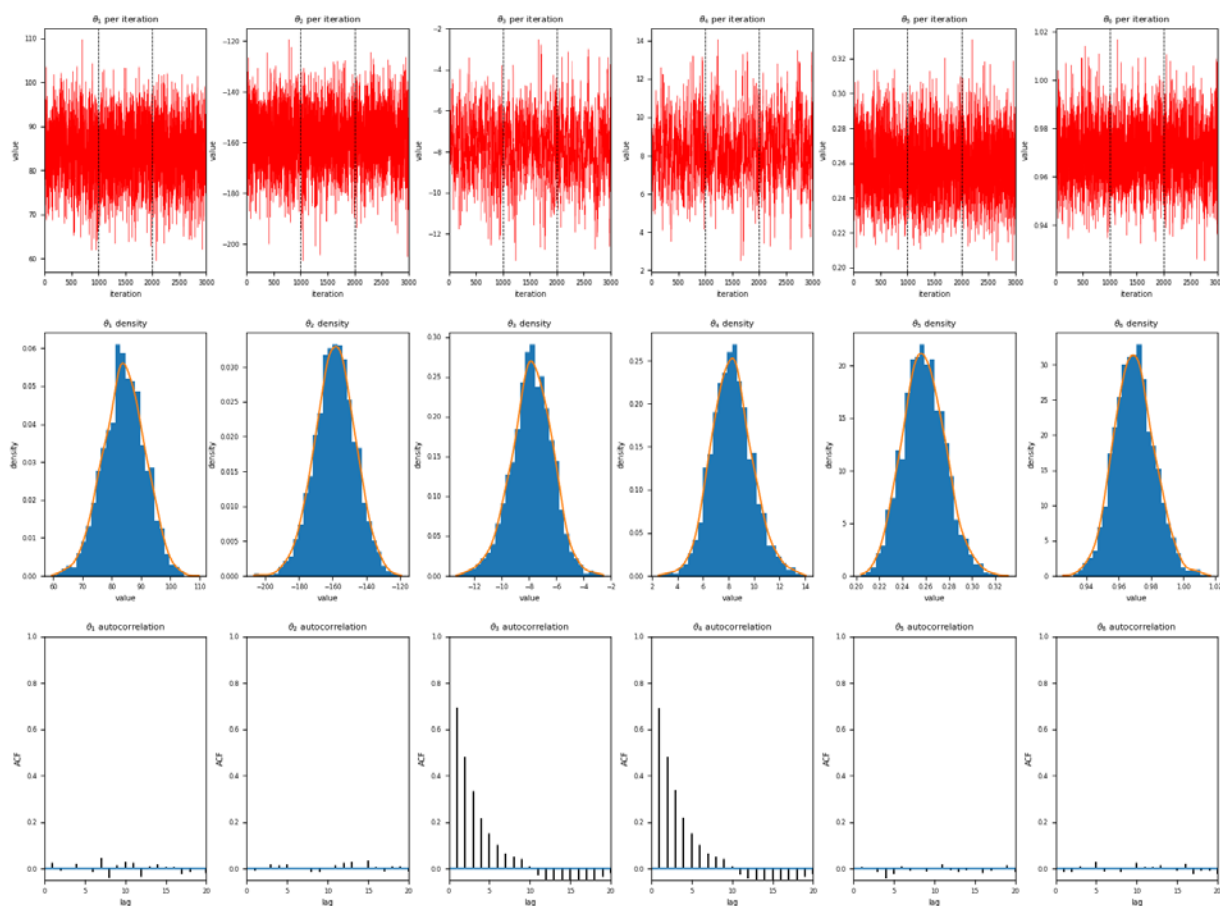
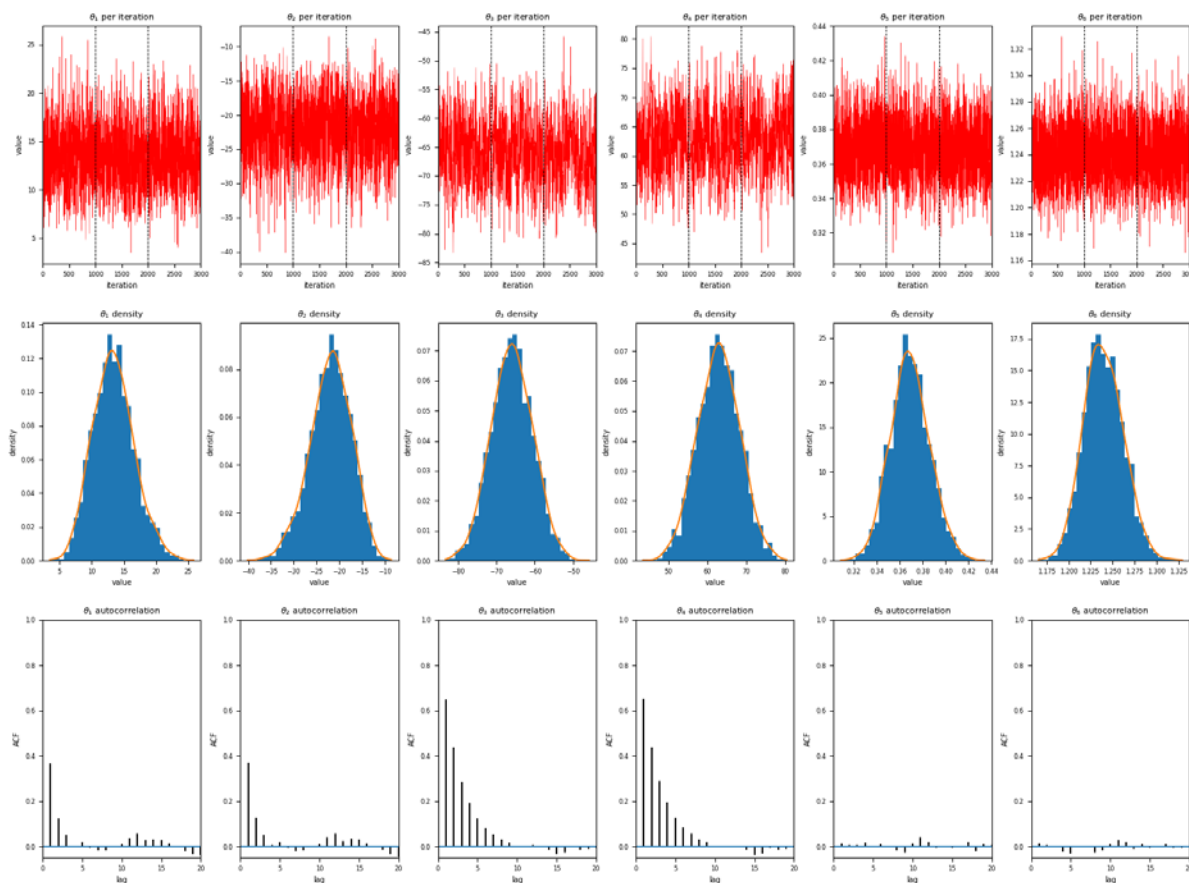


Figure D1: Values per iteration, density plot and autocorrelation of $\theta_1, \dots, \theta_6$ for low-quality buildings. The dashed lines in the top row represent the cut-offs between the 3 chains used in Gibbs sampling.



330 **Figure D2:** Values per iteration, density plot and autocorrelation of $\theta_1, \dots, \theta_6$ for medium-quality buildings. The dashed lines in the top row represent the cut-offs between the 3 chains used in Gibbs sampling.

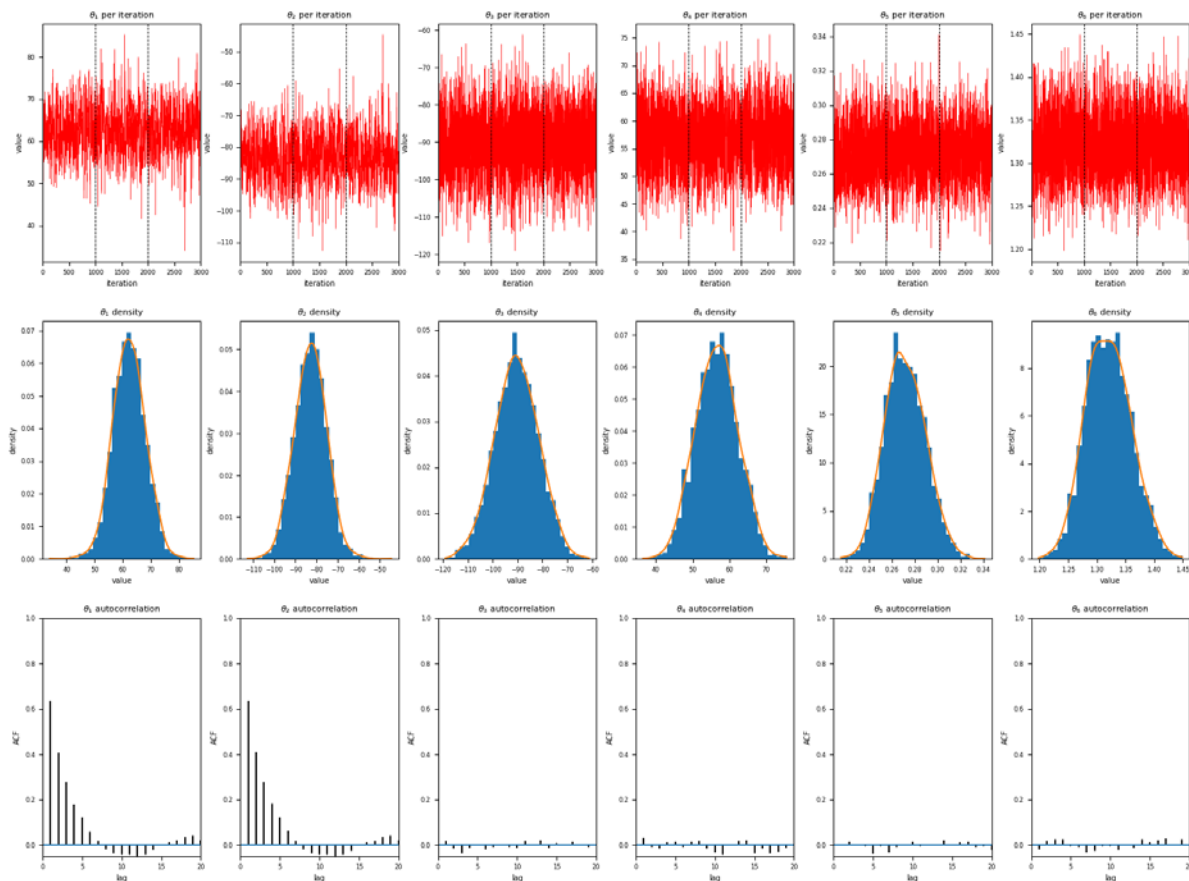


Figure D3: Values per iteration, density plot and autocorrelation of $\theta_1, \dots, \theta_6$ for high-quality buildings. The dashed lines in the top row represent the cut-offs between the 3 chains used in Gibbs sampling.

335 6 Code availability

Python and R code is available on GitHub (<https://github.com/jensdebruijn/Bayesian-updating-of-hurricane-vulnerability-functions>).



7 Data availability

340 The damage observations and wind field are available on GitHub (<https://github.com/jensdebruijn/Bayesian-updating-of-hurricane-vulnerability-functions>).

8 Author contribution

345 JB developed the methodology and took lead in writing the manuscript. JD took part in the creative process and assessed the damage ratios. AP, RG and JM collected and analyzed the exposure data. MR assisted in writing the manuscript. SK and HM assisted in development of the methodology. NB created the wind field. JA assisted in writing and oversaw the creative process.

9 Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

10 Disclaimer

350 The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of its Executive Directors or the governments they represent.

11 Acknowledgements

355 Our research was funded by an NWO-Vici grant from the Netherlands Organisation for Scientific Research (NWO; grant number 453-14-006) and an EU-ENHANCE grant from the European Community's Seventh Framework Programme (FP7; grant number 308438). This work was conducted by the World Bank's Disaster Resilience Analytics and Solutions (D-RAS) Knowledge Silo Breaker (KSB), under the Global Practice for Urban, Disaster Risk Management, Resilience and Land (GPURL). This research has been funded by the World Bank and the Global Facility for Disaster Reduction and Recovery (GFDRR) grant (TF0B1140).

360

12 References

Avila, L. A., Stewart, S. R., Berg, R. and Hagen, A. B.: National Hurricane Center Tropical Cyclone Report - Hurricane



- Dorian, Miami., 2020.
- Bilham, R.: The seismic future of cities, *Bull. Earthq. Eng.*, 7(4), 839–887, doi:10.1007/s10518-009-9147-0, 2009.
- 365 Bono, F. and Gutiérrez, E.: A network-based analysis of the impact of structural damage on urban accessibility following a disaster: The case of the seismically damaged Port Au Prince and Carrefour urban road networks, *J. Transp. Geogr.*, doi:10.1016/j.jtrangeo.2011.08.002, 2011.
- Bouwer, L. M.: Have Disaster Losses Increased Due to Anthropogenic Climate Change?, *Bull. Am. Meteorol. Soc.*, 92(1), 39–46, doi:10.1175/2010BAMS3092.1, 2011.
- 370 Brooks, S. P. and Roberts, G. O.: Convergence assessment techniques for Markov chain Monte Carlo, *Stat. Comput.*, 8(4), 319–335, doi:10.1023/A:1008820505350, 1998.
- de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J. and Aerts, J. C. J. H.: A Global Database of Historic and Real-time Flood Events based on Social Media, *Sci. data*, 6(1), 1–12, 2019.
- Cho, H., Rivera-Sánchez, M. and Lim, S. S.: A multinational study on online privacy: global concerns and local responses, 375 *New Media Soc.*, 11(3), 395–416, doi:10.1177/1461444808101618, 2009.
- Chung Yau, S., Lin, N. and Vanmarcke, E.: Hurricane damage and loss estimation using an integrated vulnerability model, *Nat. Hazards Rev.*, 12(4), 184–189, doi:10.1061/(ASCE)NH.1527-6996.0000035, 2011.
- Cowles, M. K. and Carlin, B. P.: Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review, *J. Am. Stat. Assoc.*, 91(434), 883–904, doi:10.1080/01621459.1996.10476956, 1996.
- 380 Cutter, S. L., Ismail-Zadeh, A., Alcántara-Ayala, I., Altan, O., Baker, D. N., Briceño, S., Gupta, H., Holloway, A., Johnston, D., McBean, G. A., Ogawa, Y., Paton, D., Porio, E., Silbereisen, R. K., Takeuchi, K., Valsecchi, G. B., Vogel, C. and Wu, G.: Global risks: Pool knowledge to stem losses from disasters, *Nature*, 522(7556), 277–279, doi:10.1038/522277a, 2015.
- Department of Statistics: 2000 Census of Population and Housing, Nassau, Bahamas., 2002.
- Department of Statistics: 2010 Census of Population and Housing, Nassau, Bahamas., 2012.
- 385 Department of Statistics: Population Projections 2010–2040, Nassau, Bahamas., 2015.
- Department of Statistics: Report of the 2013 Household Expenditure Survey, Nassau, Bahamas., 2016.
- Department of Statistics: Building Construction Statistics Report, Nassau, Bahamas., 2018.
- Desai, B., Maskrey, A., Peduzzi, P., De Bono, A. and Herold, C.: Making development sustainable: the future of disaster risk management, global assessment report on disaster risk reduction, 2015.
- 390 Douglas, J.: Physical vulnerability modelling in natural hazard risk assessment, *Nat. Hazards Earth Syst. Sci. Phys.*, 7(7), 283–288, 2007.
- Duggan, M. and Brenner, J.: The demographics of social media users, 2012, Pew Research Center’s Internet & American Life Project Washington, DC., 2013.
- Ellingwood, B. R., Rosowsky, D. V., Li, Y. and Kim, J. H.: Fragility assessment of light-frame wood construction subjected 395 to wind and earthquake hazards, *J. Struct. Eng.*, 130(12), 1921–1930, 2004.
- Ferrari, S. L. P. and Cribari-Neto, F.: Beta regression for modelling rates and proportions, *J. Appl. Stat.*, 31(7), 799–815,



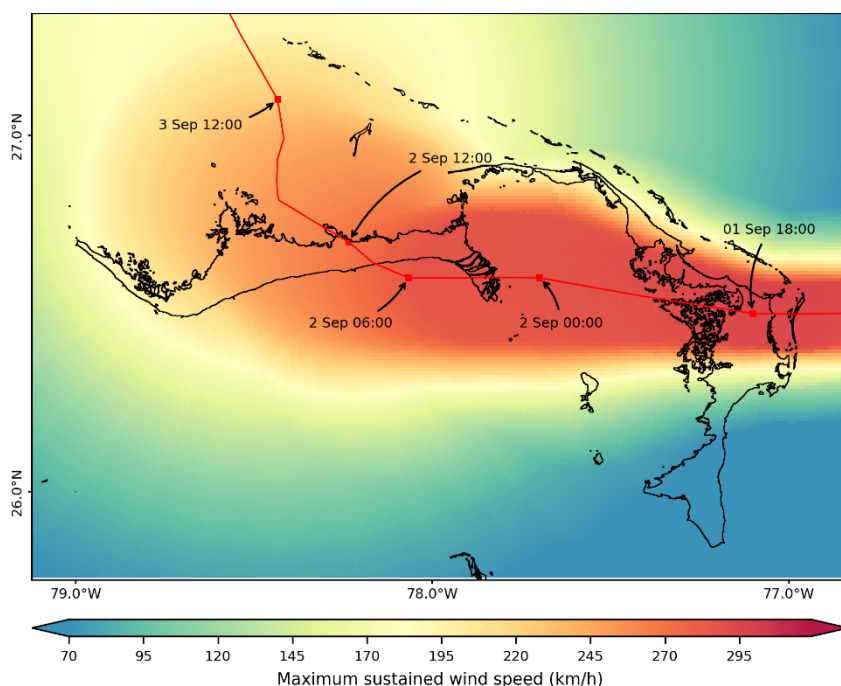
- doi:10.1080/0266476042000214501, 2004.
- Gilks, W. R., Richardson, S. and Spiegelhalter, D. J.: Markov Chain Monte Carlo in Practice Chapman & Hall: London, Markov Chain Monte Carlo Pract. Chapman Hall, London., 1996.
- 400 Gunasekera, R., Daniell, J., Pomonis, A., Arias, R. A. D., Ishizawa, O. and Stone, H.: Methodology Note on the Global RAPid post-disaster Damage Estimation (GRADE) approach., 2018.
- Gupta, A. K. and Nadarajah, S.: Handbook of beta distribution and its applications, CRC press., 2004.
- Hatzikyriakou, A., Lin, N., Gong, J., Xian, S., Hu, X. and Kennedy, A.: Component-based vulnerability analysis for residential structures subjected to storm surge impact from Hurricane Sandy, Nat. Hazards Rev., 17(1), 5015005, 2016.
- 405 Holland, G. J.: An Analytic Model of the Wind and Pressure Profiles in Hurricanes, Mon. Weather Rev., 108(8), 1212–1218, doi:10.1175/1520-0493(1980)108<1212:AAMOTW>2.0.CO;2, 1980.
- Holmes, J.: Vulnerability curves for buildings in tropical-cyclone regions, in Probabilistic Mechanics & Structural Reliability, pp. 78–81., 1996.
- Keote, S. A., Kumar, D. and Singh, R.: Construction of low rise buildings in cyclone prone areas and modification of
- 410 cyclone, J. Energy Power Sources, 2, 247–255, 2015.
- Kim, K. and Davidson, J.: Unmanned aircraft systems used for disaster management, Transp. Res. Rec., 2532, 83–90, doi:10.3141/2532-10, 2015.
- Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J. and Neumann, C. J.: The International Best Track Archive for Climate Stewardship (IBTrACS), Bull. Am. Meteorol. Soc., 91(3), 363–376, doi:10.1175/2009BAMS2755.1, 2010.
- 415 Koshimura, S., Oie, T., Yanagisawa, H. and Imamura, F.: Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from banda aceh, Indonesia, Coast. Eng. J., 51(3), 243–273, doi:10.1142/S05785634090002004, 2009.
- Koutsourelakis, P. S.: Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework, Probabilistic Eng. Mech., 25(1), 49–60, doi:10.1016/j.proengmech.2009.05.005, 2010.
- 420 Kryvasheyev, Y., Chen, H., Obradovich, N., Moro, E., Van Hentenryck, P., Fowler, J. and Cebrian, M.: Rapid assessment of disaster damage using social media activity, Sci. Adv., 2(3), e1500779, doi:10.1126/sciadv.1500779, 2016.
- Lee, K. H. and Rosowsky, D. V: Fragility assessment for roof sheathing failure in high wind regions, Eng. Struct., 27(6), 857–868, 2005.
- Li, J., Spencer, B. F. and Elnashai, A. S.: Bayesian updating of fragility functions using hybrid simulation, J. Struct. Eng. (United States), 139(7), 1160–1171, doi:10.1061/(ASCE)ST.1943-541X.0000685, 2013.
- 425 Li, Y. and Ellingwood, B. R.: Hurricane damage to residential construction in the US: Importance of uncertainty modeling in risk assessment, Eng. Struct., 28(7), 1009–1018, doi:https://doi.org/10.1016/j.engstruct.2005.11.005, 2006.
- Lin, N. and Chavas, D.: On hurricane parametric wind and applications in storm surge modeling, J. Geophys. Res. Atmos., 117(D9), doi:10.1029/2011JD0017126, 2012.
- 430 Massarra, C. C., Friedland, C. J., Marx, B. D. and Dietrich, J. C.: Predictive multi-hazard hurricane data-based fragility



- model for residential homes, *Coast. Eng.*, 151, 10–21, doi:https://doi.org/10.1016/j.coastaleng.2019.04.008, 2019.
- Mendelsohn, R., Emanuel, K., Chonabayashi, S. and Bakkensen, L.: The impact of climate change on global tropical cyclone damage, *Nat. Clim. Chang.*, 2(3), 205–209, doi:10.1038/nclimate1357, 2012.
- Miles, B. and Morse, S.: The role of news media in natural disaster risk and recovery, *Ecol. Econ.*, 63(2–3), 365–373, doi:10.1016/j.ecolecon.2006.08.007, 2007.
- Ministry of Works & Utilities: Bahamas Building Code Third Edition, Nassau, Bahamas., 2003.
- Mishra, S., Vanli, O. A., Alduse, B. P. and Jung, S.: Hurricane loss estimation in wood-frame buildings using Bayesian model updating: Assessing uncertainty in fragility and reliability analyses, *Eng. Struct.*, doi:10.1016/j.engstruct.2016.12.063, 2017.
- de Moel, H., van Vliet, M. and Aerts, J. C. J. H.: Evaluating the effect of flood damage-reducing measures: a case study of the unembanked area of Rotterdam, the Netherlands, *Reg. Environ. Chang.*, 1–14, doi:10.1007/s10113-013-0420-z, 2013.
- Mora, C., Spirandelli, D., Franklin, E. C., Lynham, J., Kantar, M. B., Miles, W., Smith, C. Z., Freel, K., Moy, J., Louis, L. V., Barba, E. W., Bettinger, K., Frazier, A. G., Colburn, J. F., Hanasaki, N., Hawkins, E., Hirabayashi, Y., Knorr, W., Little, C. M., Emanuel, K., Sheffield, J., Patz, J. A. and Hunter, C. L.: Broad threat to humanity from cumulative climate hazards intensified by greenhouse gas emissions, *Nat. Clim. Chang.*, 8(12), 1062, doi:10.1038/s41558-018-0315-6, 2018.
- Ospina, R. and Ferrari, S. L. P.: Inflated beta distributions, *Stat. Pap.*, 51(1), 111–126, doi:10.1007/s00362-008-0125-4, 2010.
- Pita, G., Pinelli, J. P., Gurley, K. and Mitrani-Reiser, J.: State of the art of hurricane vulnerability estimation methods: A review, *Nat. Hazards Rev.*, 16(2), 1–16, doi:10.1061/(ASCE)NH.1527-6996.0000153, 2015.
- Pita, G. L., Pinelli, J. P., Gurley, K. R. and Hamid, S.: Hurricane vulnerability modeling: Development and future trends, *J. Wind Eng. Ind. Aerodyn.*, 114, 96–105, doi:10.1016/j.jweia.2012.12.004, 2013.
- Plummer, M.: DSC 2003 Working Papers JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, [online] Available from: <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/>, 2003a.
- Plummer, M.: JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, in *Proceedings of the 3rd international workshop on distributed statistical computing*, vol. 124, pp. 1–10., 2003b.
- Powell, M., Soukup, G., Cocke, S., Gulati, S., Morisseau-Leroy, N., Hamid, S., Dorst, N. and Axe, L.: State of Florida hurricane loss projection model: Atmospheric science component, *J. Wind Eng. Ind. Aerodyn.*, 93(8), 651–674, doi:https://doi.org/10.1016/j.jweia.2005.05.008, 2005.
- Rossetto, T., Ioannou, I. and Grant, D. N.: Existing Empirical Fragility and Vulnerability Relationships: Compendium and Guide for Selection, *GEM Tech. Rep.*, 01(October), 77, doi:10.13117/GEM.VULN-MOD.TR2015.01, 2015.
- Shanty Town Task Force: Abaco Shanty Town Assessment Report, 2018, Nassau, Bahamas., 2018.
- Shultz, J. M., Russell, J. and Espinel, Z.: Epidemiology of Tropical Cyclones: The Dynamics of Disaster, Disease, and Development, *Epidemiol. Rev.*, 27(1), 21–35, doi:10.1093/epirev/mxi011, 2005.
- UNDP, World Bank, E.: Post-Disaster Needs Assessments., 2013.



- 465 UNISDR: Global assessment report on disaster risk reduction 2015. [online] Available from:
<https://www.unisdr.org/we/inform/publications/42809>, 2015.
- UNISDR: Report of the open-ended intergovernmental expert working group on indicators and terminology relating to
 disaster risk reduction, 2016.
- Vickery, P. J., Skerlj, P. F., Lin, J., Twisdale, L. A., Young, M. A. and Lavelle, F. M.: HAZUS-MH hurricane model
 470 methodology. II: Damage and loss estimation, Nat. Hazards Rev., 7(2), 94–103, doi:10.1061/(ASCE)1527-
 6988(2006)7:2(94), 2006.
- Wijayanti, P., Zhu, X., Hellegers, P., Budiyono, Y. and van Ierland, E. C.: Estimation of river flood damages in Jakarta,
 Indonesia, Nat. Hazards, 86(3), 1059–1079, doi:10.1007/s11069-016-2730-1, 2017.
- Winsemius, H. C., Aerts, J. C. J. H., Van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., Kwadijk, J. C. J.,
 475 Ligtoet, W., Lucas, P. L., Van Vuuren, D. P. and Ward, P. J.: Global drivers of future river flood risk, Nat. Clim. Chang.,
 6(4), 381–385, doi:10.1038/nclimate2893, 2016.
- Yu, M., Yang, C. and Li, Y.: Big Data in Natural Disaster Management: A Review, Geosciences, 8(5),
 doi:10.3390/geosciences8050165, 2018.



480

Figure 1: Maximum 1-minute average sustained wind speeds at 10 meter above surface level at 0.01° resolution for Hurricane Dorian during its passage over the Bahamas in September 2019. Time stamps are given in UTC.

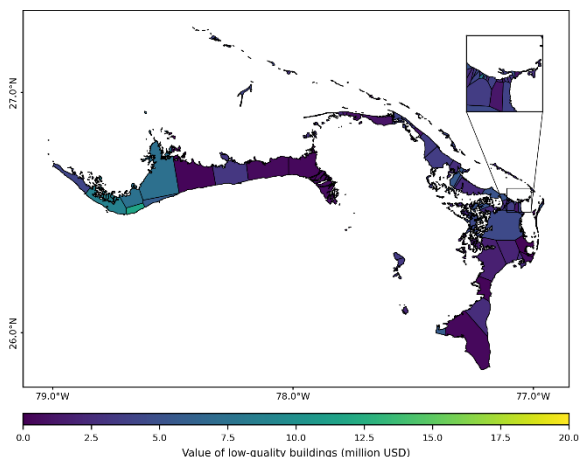


Figure 2: Value of low-quality buildings per region.



Figure 3: Footage from “Dorian’s Destruction of Treasure Cay, Abaco Bahamas” (<https://www.youtube.com/watch?v=hvCQotLWW-y4>) containing (left) a medium quality building with structural detailing with roof damage and water intrusion to part of the building’s structure, non-structural damage (ca. 25% damage) and (right) a low quality building (with minor structural detailing), missing its front wall, roof damage, and significant debris (ca. 50% damage).

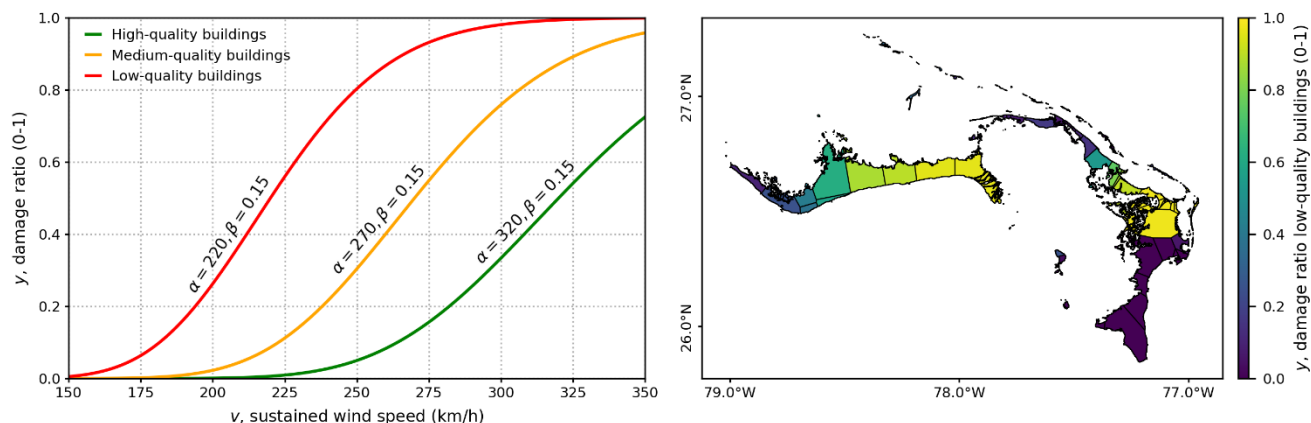


Figure 4: Vulnerability functions for low-, medium- and high-quality buildings (left) and damage ratios for low-quality buildings in Grand Bahama and the Abaco Islands (right).

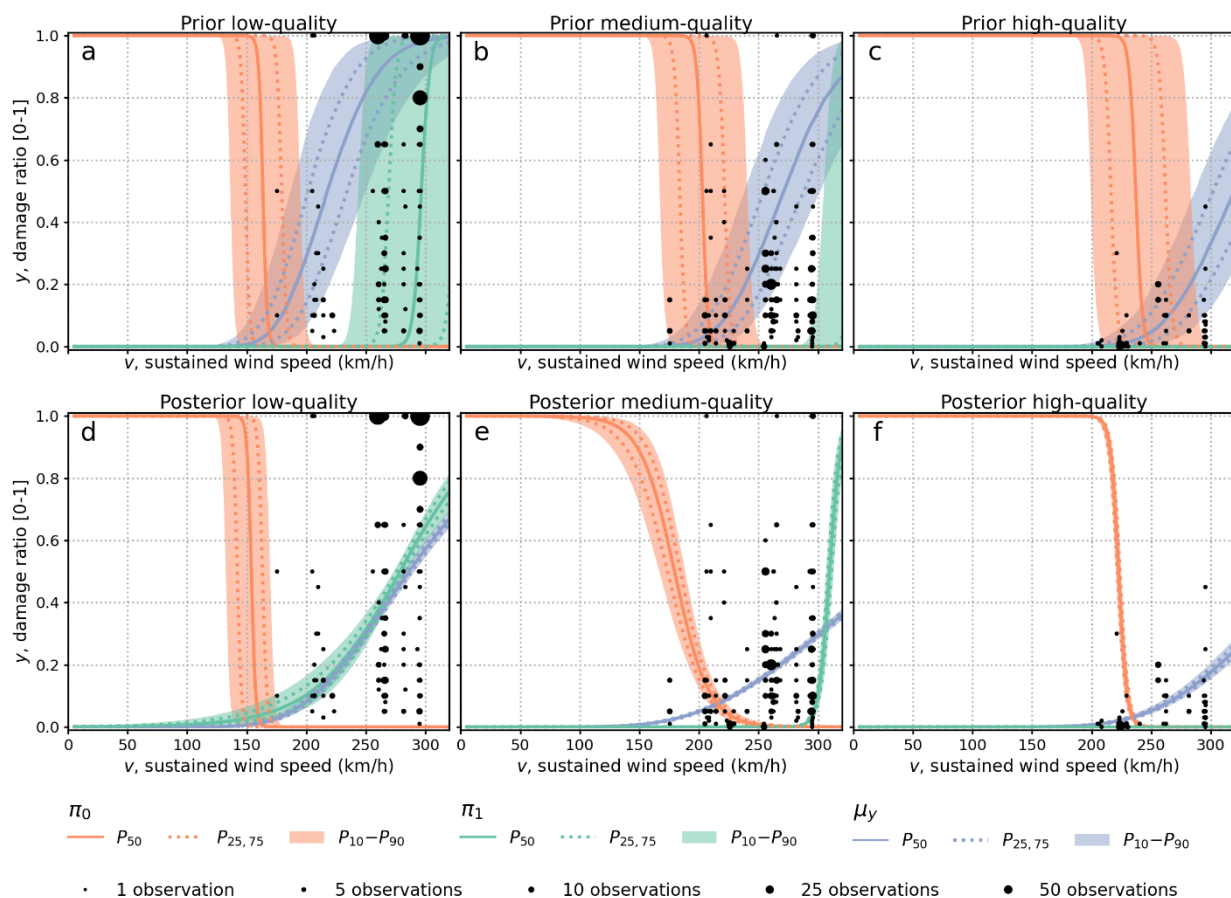


Figure 5: Visualization of observations, priors and posteriors 25th, 50th (median), 75th percentile and the 10th–90th percentile range for low-, medium- and high-quality buildings.

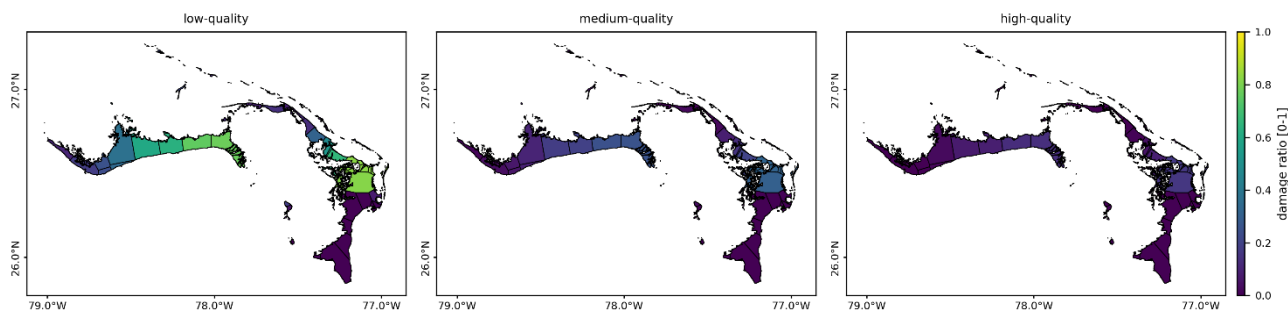


Figure 6: Posterior damage ratio per district for low-, medium- and high-quality buildings.

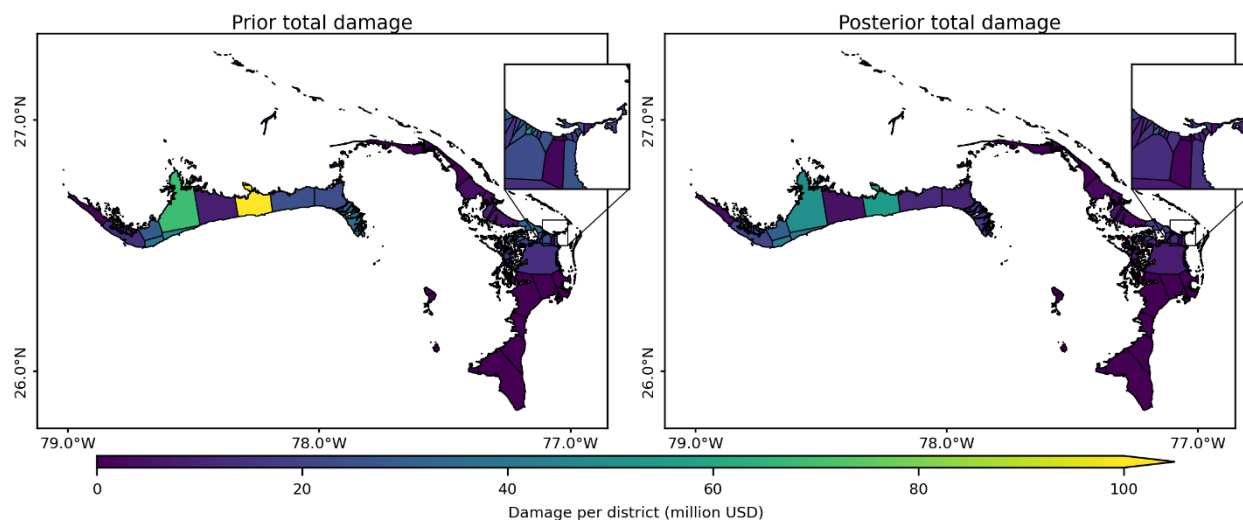


Figure 7: Damage to low-, medium- and high-quality housing combined derived from prior and posterior vulnerability curves.

500 **Table 1: Damage for residential buildings per building class derived from prior and posterior distributions**

	Prior (million USD)	Posterior (million USD)	Percentage change
<i>Low-quality</i>	155	124	-20%
<i>Medium-quality</i>	543	324	-40%
<i>High-quality</i>	359	210	-41%
Total	1056	658	-38%