Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2020-282-AC1, 2021

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Using rapid damage observations from social media for Bayesian updating of hurricane vulnerability functions: A case study of Hurricane Dorian" *by* Jens A. de Bruijn et al.

Jens A. de Bruijn et al.

jens.de.bruijn@vu.nl

Received and published: 25 January 2021

The paper builds on important work performed on estimating damages from imagery. However, the current literature review and methodology sections are cursory at best and lack significant details for using images in a damage assessment. For instance, the literature review is missing any detail on quantifying structural damages through survey such as using the Tornado Injury Scale (TIS; Curtis & Fagan, 2013), or papers such as Meyer and Hendricks (2018) which directly measure damages and recovery using images. Like the literature review,

the methods section was limited in terms of image estimates. The authors do a great job outlining the model they use for the final analysis, but their contribution is using youtube to inform those models. The authors allocate two pages to explaining the Bayesian model, but they spend two paragraphs explaining how they collected the imagery. This lack of detail limits the use of this paper for other researchers and removes any valid reproducibility.

First of all, we thank the reviewer for your kind words and suggestions. Below, I will list the reviewers points in order and discuss their suggestions. Here, we also specify specifically how we aim to make the contribution more reproducible.

Several questions I asked while reading the methods included:1) how many people watched the videos and quantified damages? 2) If more than one person was watching the videos and estimating damages how did the authors deal with potential issues with inter-rater reliability (See meyer and Hendricks for example)? 3) How did you rate damages? Without answers to questions like these the generalizability of the study is severely limited.

Thank you for your suggestion. We will in more detail discuss literature on quantifying structural damages through images, for example by including a map of the observations and their source (i.e., observation from ground or air).

Unfortunately, the scales suggested by Curtis & Fagan (2013) and Meyer and Hendricks (2018) use a damage score (TIS1-10 and 0-9 respectively), but do not translate these scores into a damage ratio.

However, we fully agree with the reviewer that this leads to problems regarding reproducibility, especially because the damage was assessed by only one person (James E. Daniell) leading to subjectivity in the scores. In fact, this is a common problem when human judges are used. A solution to this problem is to use multiple judges assuming that the average or median assessment of multiple judges will lead to a more objective, and reproducible, score.

Therefore, to obtain a more objective score Antonios Pomonis and Joshua Macabuag assessed the damages and building classes in each image as well resulting in three damage classifications for each image by engineers experienced in assessing building damages after disasters. Then, following Meyer and Hendricks (2018) and others we calculated intercoder reliability tests. Subsequently, we will use the median damage ratio and building class for further analysis.

From a preliminary analysis of the scores for the individual buildings with three judges, where each judge rates each target, we obtain an intraclass correlation of 0.92 for building damages using the Spearman Brown adjusted reliability and a Fleiss kappa of 0.30 for building class.

Within the results section I was disappointed to not see a section on how these measures were validated. I don't feel it is enough to say "total damages are lower with this new model", without first giving evidence as to how your estimates improved the calculations. Without these validation metrics from a test dataset, the findings can't be assumed to improve the model and may in fact be making it worse.

We fully agree with the reviewer that a comparison with gold standard vulnerability curves would be beneficial. A (cursory) damage report is available for hurricane Dorian (ECLAC et al., 2019). However, this damage report only reports total damages. This means that 1) the report also includes damage from the storm surge, and 2) no vulnerability curves are presented hindering a direct comparison of vulnerability curves.

To make a true comparison, between our vulnerability curve and a gold standard vulnerability curve, we would require not just information about risk (the damages), but rather about the vulnerability component of risk. However, here we run into two problems:

1. To the best of our knowledge no independent wind vulnerability curves are avail-

able for hurricane Dorian in the Bahamas, while comparison with similar events in similar locations would neglect the purpose of this manuscript of creating event-specific vulnerability curves.

2. Even if vulnerability curves were available, these are dependent on both the hazard component and damage observations. Since wind speeds as part of the uncertainty within the process, many vulnerability functions being characterized as semi-empirical (e.g., Mason and Parackal, 2015; Pita et al., 2015; Smith et al., 2020; Walker, 2011) and direct comparison would be unproductive.

Moreover, other vulnerability curves and observations would also be prone to uncertainties. Therefore, we believe that by updating previously existing evidence with new data, we are in fact converging towards the true vulnerability curve. Any additional observations should be treated as additional evidence rather than test data.

However, we should make this clearer in the manuscript and if we were allowed to submit a revised version of our manuscript, we shall include a discussion to this extent. In addition, we will revise several sentences. For example, *"total damages are lower with this new model"* as quoted by the reviewer to *"using the posterior vulnerability curves total damages are projected to be lower"*.

References

Curtis, A. and Fagan, W. F.: Capturing damage assessment with a spatial video: An example of a building and street-scale analysis of tornado-related mortality in Joplin, Missouri, 2011, Ann. Assoc. Am. Geogr., 103(6), 1522–1538, 2013.

ECLAC, IDB, PAHO and WHO: Assessment of the Effects and Impacts of Hurricane Dorian in the Bahamas., 2019.

Mason, M. S. and Parackal, K. I.: Vulnerability of buildings and civil infrastructure to tropical cyclones: a preliminary review of modelling approaches and literature, Bush.

Nat. Hazards CRC, Melb., 2015.

Meyer, M. A. and Hendricks, M. D.: Using photography to assess housing damage and rebuilding progress for disaster recovery planning, J. Am. Plan. Assoc., 84(2), 127–144, 2018.

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.

Smith, D. J., Edwards, M., Parackal, K., Ginger, J., Henderson, D., Ryu, H. and Wehner, M.: Modelling vulnerability of Australian housing to severe wind events: past and present, Aust. J. Struct. Eng., 21(3), 175–192, 2020.

Walker, G. R.: Modelling the vulnerability of buildings to wind—a review, Can. J. Civ. Eng., 38(9), 1031–1039, 2011.

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2020-282, 2020.