**Note:**

The text in black is reviewers’ comments, blue text is authors’ response to reviewers, and highlighted text is the changes in manuscript and is also copied here to give reviewers easy access to the changes.

**Reviewer 1**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Yes | Can be improved | Must be improved | Not applicable |
| Does the introduction provide sufficient background and include all relevant references? | ( ) | ( ) | (x) | ( ) |
| Is the research design appropriate? | ( ) | (x) | ( ) | ( ) |
| Are the methods adequately described? | ( ) | (x) | () | ( ) |
| Are the results clearly presented? | ( ) | (x) | () | ( ) |
| Are the conclusions supported by the results? | ( ) | (x) | () | ( ) |

Comments and Suggestions for Authors

This is an interesting paper that draws on unique data and links together computing, communications, and social sciences. The paper is fairly well written and would be informative for the journal readership. I like how the paragraphs are concise and the logic of the paper is well thought out.

I do have four worries - that would be easily amended with some thought and inclusion of a few more sentences in the lit review and discussions and conclusions:

1. This paper puts forth that notion that citizen science derived information is inaccurate until verified against 'scientifically-derived and confirmed' data. There has been quite a shift in that kind of positivistic school to one of participatory and other approaches to understanding the world that emphasized the complementarity of different ways of knowing (see for example, Williams, L. (2017). Empowerment and the ecological determinants of health: three critical capacities for practitioners. *Health promotion international*, *32*(4), 711-722.). **Can the authors speak to why they assumed this sort of hierarchical relationship?** If warranted, that's fine, but they need to specify this positionality at the outset of the paper, otherwise, they need to speak to this relationality and paradigm shift - and frame the papers' use of datasets generated from western and citizen science as advantageous in gaining a fullsome understansing of societies experiences of flooding and other risk communications.

Thanks for the suggestion! Yes, there has been a paradigm shift in resilience and risk communication that is more participatory. However, citizen science and crowdsourcing are two different participatory approaches used for scientific problem solving. Twitter is a social media site that depends on crowdsourcing. A discussion of the difference between crowdsourcing and citizen science, and the reasoning behind out use of a hierarchical relationship is presented in the introduction section to set the stage for this research. The section is highlighted below.

**(Line 45-56)** Citizen science-based platforms (e.g., iCoast, Tweet Earthquake Dispatch, CitizenScience.gov) allows citizens to collaborate with scientists in collecting and analyzing data, reporting observations and disseminating results about scientific problems (Bonney 1996). Crowdsourcing platforms, such as Twitter and Facebook, are social media and social networking sites, that allow non-experts to generate new knowledge and data sets(Greengard 2011, Hetmank 2013). Although, both citizen science and crowdsourcing engage socio-culturally diverse and geographically dispersed citizens for data and knowledge creation/collection, each has subtle differences(Onsrud, Camara et al. 2004, Onsrud and Campbell 2007). While crowdsourcing remains an ill-defined approach that uses large networks of people, citizen science solely uses scientists, volunteers, and lay people with interests and knowledge about a specific topic (Wiggins and Crowston 2011). Because tweets are generated via crowdsourcing and tend to contain rumors and hoaxes, we assumed the tweets to be inaccurate and implemented a hierarchical approach to verify the reliability and relevant of the tweets using scientifically derived and confirmed data.

2. What are some of the social psychological concepts that drive people to spread information via social media - and get their information there? What biases and heuristics do people employ when seeking and sharing risk communications? The paper is very technical in nature but really leaves out the human dimensions. That part can be improved.

Thank you for the great comments, we added the following highlighted sections addressing your concerns about missing human dimensions:

**(Lines 40-42)** From social psychological perspective, reasons that drive people to share information on social media include but not limit by self-efficacy, self-fulfilment, altruism, social engagement, reciprocity, and reputation (Moon 2014, Oh and Syn 2015).

**(Lines 63-67)** Finally, heuristic plays a big role in deciding what or whether to share information on social media and this becomes more influential in complicated and unanticipated crisis situations, which will inevitably introduce errors and biased judgements to shared risk information (Dale 2015).

3. The discussion does not return to any of the literature presented earlier for comparison or theory critique or theory building. The question "So what?" comes to mind. There are no recommendations made for any potential stakeholders: people living in flood prone areas, risk communicators, social media managers, local, regional and national policymakers, researchers overall, or emergency responders.

We addressed these aspects in the discussion and future work section by adding the following highlighted text.

**(Lines 388-395)** Considering the limitation of this research workflow, future research should focus on streamlining the process and automating the entire workflow of assessing relevance and reliability of Twitter data. Moreover, encouragement and involvement of citizen-led reliability evaluation efforts following well- informed protocols will greatly boost the usefulness of this research workflow. While researchers are working to maximize the amount of risk information from Twitter, it is essential for emergency management agencies to develop easy-to-follow standards that are tailored for Twitter users to encourage dissemination of relevant and reliable crisis information to facilitate their use for response activities.

4. The paper's limitations are technical - there is no reflection in human error, alternative interpretation, and future work.

Thank for bringing this up. We added several sentences explaining how to minimize human error and limitation. These are included in the future work and discussion section.

4.1. Evaluation of text content

**(Line 189-190)** Three authors of this paper worked on manual evaluation of tweet text and each tweet was evaluated by at least two authors to minimize human error or bias.

5. Discussion, implications for risk communication, and future research

**(Line 386-387)** Finally, human errors and heuristic bias may be introduced in manual approaches, even though multiple authors cross-checked the results.

Details:

Section 3.2.1 - what verification was performed to assess whether the keywords you used were effective at getting an adequate sample of tweets?

Thank you for the question. Our previous publication (Liu et al., 2018) used the same set of keywords to extract 5000 tweets with varying relevant scores (derived from MongoDB’s function) and we examined the relationships between tweet volume and its content with precipitation amount, damage extent, and official reports. The result showed that the tweets with moderate to high relevant scores were highly correlated to the flooding events in terms of precipitation volume and damage extent spatiotemporally.

Did you try these words in a scoping/systematic review of academic and grey literature too? or aginst other social media channels to figure out if these words reflected the discourse?)

Before collecting Twitter data, we consulted similar research and found that it is a common practice to define a preliminary set of keywords and then recursively add more keywords based on literature search to form the final set of keywords. Though we did not check those keywords in other social media channels, but we created a set of keywords based on emergency management and risk communication literature.

Did you check on common spelling mistakes and search for tweets with spelling mistakes of those terms?

As for spelling mistakes, the natural language processing library we used would skip those tweets that did not fit our search criteria, whether due to spelling error or not.

Reference

Xiaohui Liu, Bandana Kar, Chaoyang Zhang & David M. Cochran (2018) Assessing relevance of tweets for risk communication, International Journal of Digital Earth, 12:7, 781-801, DOI: [10.1080/17538947.2018.1480670](https://doi.org/10.1080/17538947.2018.1480670)

- Line 48 remove little

Thank for the suggestion. The word “little” has been removed.

- Awkward wording lines 78-80

Thank for identifying the issue. It has been fixed (see below).

**(Line 95-96**) Despite the abundance of existing evaluation methods, some algorithm-based studies rarely incorporate potentially relevant external data sources to the research context, such as…

**References**

Bonney, R. (1996). "'Citizen Science: A Lab Tradition' [in] Living Bird: For the Study and Conservation of Birds." Living Bird: For the Study and Conservation of Birds **15**(4): 7-15.

Dale, S. (2015). "Heuristics and biases: The science of decision-making." Business Information Review **32**(2): 93-99.

Greengard, S. (2011). "Following the crowd." Communications of the ACM **54**(2): 20-22.

Hetmank, L. (2013). "Components and Functions of Crowdsourcing Systems-A Systematic Literature Review." Wirtschaftsinformatik **4**: 2013.

Moon, G. (2014). "Why People Share: The Psychology of Social Sharing." from <https://coschedule.com/blog/why-people-share/>.

Oh, S. and S. Y. Syn (2015). "Motivations for sharing information and social support in social media: A comparative analysis of Facebook, Twitter, Delicious, YouTube, and Flickr." Journal of the Association for Information Science and Technology **66**(10): 2045-2060.

Onsrud, H., G. Camara, J. Campbell and N. S. Chakravarthy (2004). Public commons of geographic data: Research and development challenges. International Conference on Geographic Information Science, Springer.

Onsrud, H. and J. Campbell (2007). "Big opportunities in access to" Small Science" data." Data Science Journal **6**: OD58-OD66.

Wiggins, A. and K. Crowston (2011). From conservation to crowdsourcing: A typology of citizen science. 2011 44th Hawaii international conference on system sciences, IEEE.