Title: Climate risks, digital media, and big data: following communication trails to investigate urban communities' resilience

Authors: Rosa Vicari, Ioulia Tchiguirinskaia, Bruno Tisserand, Daniel Schertzer

<u>Summary and overview:</u>

The aim of this study is to use web data produced during and after a flood to understand the interactions between the event and the social perception of the event using big data analysis techniques. The study looks at two different flood events in France (Alpes Maritimes 2015 and Paris 2016) and analyses media coverage of the event in French news articles and on Twitter. It demonstrates useful methods for understanding the public perception of flood events that could be helpful in improving the management of such events. This work is novel, relevant and worthwhile. However, we have suggested minor corrections mainly around including a definition of urban resilience and a few clarifications.

<u>General comments:</u>

Paragraphs should not be single sentences. The first sentence of the paragraph should set the topic of the paragraph, and the remaining sentences should develop the topic.

It would be clearer if figures that are central to the discussion are included within the article itself rather than as a supplement, if possible.

Urban resilience should be clearly defined, since the study is framed in the context of urban resilience assessment (see for example "Defining urban resilience: A review", Sara Meerow et al.). Similarly, some background on what urban resilience is should be included in the introduction. This can be used to explain why certain topics are expected to appear in the press before/during/after a flood event. Currently topics relevant to flood resilience are assumed implicitly (e.g. p.17 "it is interesting to note that there is no discussion of nature-based solutions") but there is no explicit justification for this.

<u>Specific comments:</u>

p.3: "Animals being destroyed" is not the common phrasing

p.3: A sentence introducing the Seine event in similar terms to the Alpes-Maritimes event would be useful. What was the driver of flooding, and from which date to which date was the area in fact flooded?

p.3: A different number of days before the event is used in each case, why is this? How are start and end dates for the analysis selected?

p.4: Not clear what exactly is meant by "This bond has been progressively fading". Does it mean that the power of the media to shape public opinion is decreasing?

p.4: Why is Twitter not also used for the Alpes-Maritimes event?

p.5 check grammar: toone

p.6: Is Europresse free to use?

p.7 section 3.1.5: What kind of quantitative analysis was done once the data was exported to Excel, what was the purpose of exporting the data?

p.8: Clarify "the peak of published terms per day is reduced"

p.8: Clarify "less progressively"

p.11: Is "red vigilance" is not commonly used in English, is there a better way to translate this?

p.13 section 4.3.1: It would be good to give examples of words that are in each of the clusters in the bullet points (like in section 4.2.1) especially since the figures are in French and will not be understandable by all the readers.

p.14 figure 4: Not clear how the distinction is made between terms relating to risk (which is normally defined as hazard x impact) and those relating to the weather events and impact. Also not clear what is meant by 'media content'.

p.14: Could individual accounts not also be those of professionals/individuals in positions of responsibility? In that case, how do you distinguish between individual accounts with private interests and those who speak from a professional role?

p.15 figure 5a: If I understood correctly the intention of Figure 5a), it might be clearer to show it as a stacked bar chart that sums to 100 % overall but with two separate colours, one for individual and one for company accounts (with a legend specifying which colour corresponds to which type of account)

p.16: Grammar point; "…not AS significant as"

p.17: After "an unexpected result is that some stakeholders involved in flood risk management in Paris are not visible in the media debate in 2016" please provide some examples of which stakeholders are missing.

p.18: perhaps "relation" would be more accurate than "correlation" to explain the comparison that is made between river discharge and the number of publications on the topic.

p.18 line 24: It might be clearer to add in brackets the two locations in to clarify that two independent comparisons are made, one between the two locations, and another between two different types of media in one of the two locations.

p.18 quality of the content (not pluralised)