

Response to Reviewer 1

We would like to thank the reviewer for the very detailed and insightful comments to the first draft of the paper, which have helped to greatly improve it. We agree that many details needed further clarification. We hope that this improved version is easier to understand and makes the key points in a clearer fashion.

Comments from Reviewer 1	Response to Comments
This paper discusses a cost-benefit analysis for citizen observatories based on results in a specific catchment. The content is relevant and will be a valuable addition to citizen science research.	We thank the reviewer for their positive comment.
Unfortunately, in its current form the paper is difficult to follow and lacks information to fully understand the content. Specifically the citizen observatory need to be explained in much greater detail in the methodology chapter. What exactly does the citizen observatory measure? Are sensors being used and if so for what variables? Are observations only made during floods? How many volunteers were used? Do the volunteers get paid or is this part of their job? And in what way do these observatories reduce the cost of floods? These questions are very central to this paper and are currently not communicated.	We have added a separate section before the methodology to fully describe the citizen observatory including what is measured and information about the volunteers. The cost reduction is related to decreasing the vulnerability component of risk, which is described in more detail in the methodology.
Furthermore, for some values (e.g. the weights) no rational or reference is given.	See the responses to the specific comments below regarding the different values used in the analysis.
A full assessment of the paper can only be made once this information is provided. I therefore encourage the authors to resubmit the manuscript after including this pertinent information and some other revisions.	We understand and have taken careful note of the reviewer's comments and addressed them as detailed below in response to the more specific comments.
Specific comments	
In general, the readability of the paper is poor. One option would be to reduce the number of abbreviations. The authors should also avoid long sentences when possible. Phrases such as "detriment of an increase" (L 422) are confusing – does that mean a decrease? In addition the paper could focus more on the main story (i.e., what is written in the abstract).	We have thoroughly edited the manuscript to improve the readability.
There are many assumptions presented in the methodology. The discussion should include a paragraph that discusses the potential effects of these assumptions and the likely uncertainty due to these assumptions.	A paragraph was added to the Discussion and Conclusions section within the limitations regarding the assumptions made in the methodology, the potential effects, etc.

Comments from Reviewer 1	Response to Comments
Regarding costs (e.g. in table 12): If you are talking about costs in the millions and billions I would not write the value with a precision of one Euro. That is simply not realistic and gives a false impression of certainty. Ideally you would write ranges, or a value +/- another value.	These changes have been made.
The literature review and the introduction could have been more extensive. Are there any other citizen science projects that discuss costs vs. benefits? Or at least that discuss costs and financial benefits? How did they try to assess this and why is your method different, or why has this not been done yet?	The literature review in the introduction has been expanded regarding studies that have considered the costs and benefits associated with citizen science projects, e.g., the time invested by researchers in engaging and training citizens (Thornhill et al., 2016); to relate cost and participant performance for hydrometric observations in order to estimate the cost per observation (Davids et al., 2019); to estimate the costs as data-related costs, staff costs and other costs; and the benefits in terms of scientific benefits, public engagement benefits and the benefits of strengthened capacity of participants (Blaney et al., 2016); and to compare citizen science data and in-situ data (Goldstein et al., 2014; Hadj-Hammou et al., 2017). This approach is different because it calculates the damage costs that would be avoided if a citizen observatory was fully implemented in this basin.
L 9: In what way are citizen observatories a recent form of citizen science? (perhaps this will be resolved through a more thorough introduction to citizen observatories)	We have now specified the characteristics of citizen observatories further.
L 10: “over a period of time” is a vague statement and could also just be a day, in which case there is no difference to Blitzes. Is there a minimum required time?	Although we think the current expression was already clearly distinguishing between one off Blitzes and an extended time period in which a CO operates, we have further refined this.
L 20-21: Worldwide? In Italy?	The introduction has been rewritten so this is no longer relevant.
L 29: Dominican Republic – not Dominica	The introduction has been rewritten so this is no longer relevant.
L 42-48: The first part of this paragraph describes collaborative citizen science and not necessarily a citizen observatory. What exactly is the main difference between a CO and a collaborative citizen science project? Is it the inclusion of a public authority? In that case start with that sentence to describe the difference. Not all long-term collaborative citizen science projects are COs.	We have specified the characteristics of citizen observatories further and explained the required links to policy and the inclusion of public authorities.

Comments from Reviewer 1	Response to Comments
L 48-50: Not all of these references are related to a CO. Also, a more thorough explanation of what these individual papers have found would be nicer, rather than just a list of literature.	We have removed references to citizen science rather than citizen observatories from the list. Moreover, we have summarised the findings from these papers as suggested.
2.1 Input data: A table would provide a clearer picture of the input data. Also when is the pollutant data used? The conclusion specifies that contamination was not considered (L 477). Why is the data used for the hydraulic model (L 109) not included? Add a paragraph describing which data is used for what.	We have summarized the input data in a table. We considered the presence of punctual and widespread sources of pollutants in the vulnerability assessment, but we did not use a model to evaluate the actual propagation of contamination. We have included the data coming from the hydraulic model and included more information about the model in the Supplementary Materials. A column was added to the input data table indicating where the data sets are used in the methodology.
L 86-87: Is the risk assessed differently at a different scale than mesoscale? Why does this not hold for other scales? This sentence is confusing.	To avoid confusion, we have removed this sentence. However, the assessment of risk does differ in terms of the data used at different scales. For example, at the microscale, assessments are characterized by a high level of detail regarding the exposure so they would take individual assets at risk into account, such as buildings, vehicles or infrastructure. At the mesoscale, exposure information is usually based on aggregation, using a land use map to assess exposure, as is carried out in this paper.
Figure 1: Is there an arrow missing from the box “Depth, Velocity, Persistence”? Why do the “Land Use Map” and “People distribution map” have a different design? Also capitalization should be uniform.	This figure has been modified to better align with the text and to address the comments of the reviewer.
L 93-98 and Table 1: would fit better into 2.2.2	The original Table 1 has now been moved to the Supplementary Material (Table S1). We have also shifted text around in the methods section to improve clarity.
L 100: “must be addressed” → by whom or for what?	We have modified this to read: According to Article 6 of the 2007/60/CE Flood Directive (EU, 2007), when local authorities implement a Flood Risk Management Plan, three hazard scenarios must be considered: etc.
L 109: Which hydraulic model?	Details of the hydraulic and hydrological model have been added to the Supplementary Material.
Table 2: What does the sign “%” mean? Is this meant to be a “-”? (same in table 3)	This was meant to be a “-” sign and has been replaced in both tables.
L 140-142: “Greater awareness tends to correspond to greater preparation if an event takes place.” Please add a source for this statement.	This sentence has been modified as follows: Some studies have found that if citizens have directly experienced a flood, their perception of flood

Comments from Reviewer 1	Response to Comments
	risk is higher (e.g., Thistlethwaite et al., 2018) although the factors that determine flood risk perception are varied. Moreover, the results from different studies can be ambiguous and/or contradictory (Lechowska, 2018).
Figure 3: Where do the weights come from? Are they assumed by the author or based on literature? What is the reasoning behind the weights? Why is “Insurance density” stated, when the weight is 0? What is the uncertainty for these weights and how do they affect the results?	We have added the following text: These values have been developed by the Provincia Autonoma di Trento (2006) from decades of experience with understanding exposure related to flood risk. Moreover, they have been tested and shown to be valid within AAWA. The insurance density is one of the components of risk spread in the guidelines from ISPRA (2012) and appears in Figure 3 for completeness. However, in the case of the Brenta-Bacchiglione catchment, the insurance density is 0.0 because insurance companies do not currently offer premiums to protect goods against flood damage.
Figure 4: What is the normalized index function and how was this value computed? What does it mean?	This is now Figure S1 in the Supplementary Material. This has been referred to as a value function to be consistent throughout the text. We also explain what value functions are.
Figures 4-6 could also be included in a supplementary material. In general section 2.2 is rather long and according to the abstract not central to this paper. Parts of this section could be included in a supplementary material.	Figures 4 to 6 have been added to the Supplementary Material.
L 217: You could add that people are also less likely to take warnings serious in future.	This statement has been added to the relevant section of the paper.
L 242: Why are road inundations lumped together, regardless of whether or not the infrastructure is damaged? Surely the cost differs significantly in those two scenarios?	Here the vulnerability is related to whether the road network can be used or not, i.e., at what stage cars become unstable during a flood. Hence, the vulnerability values are based on the estimation of the critical water height and velocity for the stability of vehicles during a flood (and have nothing to do with damage to the network infrastructure). The data have been derived from direct observation in laboratory experiments and from a report on the literature in this area (Reiter, 2000; Shand et al., 2011). In the estimation of the direct costs of the flood, the costs for damage to buildings, soil and infrastructure (including roads) assumes complete rebuilding or restoration and therefore is dealt with separately.

Comments from Reviewer 1	Response to Comments
L 256: Does this mean that the vulnerability depends on the seasons, i.e., on when the crops are growing?	No, this refers to Corine Land Class 9: Unproductive land, which is defined as unproductive for the whole year.
L 260-266: I do not understand why “contamination/ pollution, erosion” and “open space” are lumped together here? Also apart from describing what an open space is, the connection to flood susceptibility unclear. Please make the connections clearer to the reader.	We have modified this section and focussed on contamination/pollution and erosion.
Why is the vulnerability high (i.e., 1) when “Integrated Pollution Prevention and Control (IPPC) installations” are present? Shouldn’t these installations reduce the vulnerability?	This has been modified as follows: ...the presence of relevant pollution sources was identified (Tables 1 and S1) and assigned a vulnerability of 1.
L 273-274: Please rephrase. I suppose “cultural heritage” is not actually considered an “adverse consequences of future flood events”.	We have removed this phrase.
Figure 9: Why does water height not play a factor in (d)?	From experimentation (Citeau, 2003), the height of the water does not affect the vulnerability of natural and semi-natural environments up to a water height of 2.5m.
L 284: Repeat what the macro-categories are.	These macro-categories have been repeated although we have also modified these sections.
L 287: I would rephrase “moderate” as “low”, as moderate is actually a synonym of medium. (same in Table 5) You actually use “low” already in Table 9 – please homogenize.	This has been modified in the text and in Table 5.
L 301: This is the first time you mention a reward. What reward do you mean? “Direct and indirect users” -> of what?	We have modified this section so this is no longer relevant.
L 308: One variable should not have multiple letters in an equation. This could be considered to be $I^*S^*R^*R$. (Also applies to some other formulas in this paper.)	We do not completely agree with this comment. For example, evapotranspiration is commonly referred to as ET, potential evapotranspiration as PET and we can find many examples in the literature (including in HESS) of multiple letter variables, e.g. RMSE, NSE, etc., which are used very regularly in hydrology.
Table 6: Where do the weights come from? Are they calculated or taken from another source?	The weights in this table have been determined through expert consultation within the Alto-Adriatico Water Authority (AAWA), supported by the guidelines from ISPRA (2012).
L 330: superscript for 2	This has been corrected.
L 349: Please define what you mean with a “unit of management”.	A unit of management is part of terminology from the INSPIRE Directive but we have removed this to avoid any confusion.

Comments from Reviewer 1	Response to Comments
L 358-359: "Because rapid floods are difficult to predict, early warning systems and prevention measures are of less use in this region." I have some doubts regarding this statement. Surely early warning systems (sirens) and prevention measures (reforestation) still reduce the risks associated with rapid floods? Please offer a source or rephrase.	We have modified this paragraph and removed this statement.
L 361: superscript for th	This has been corrected.
L 365: Was data that was not collected by trained volunteers discarded? How many volunteers were trained and how many were not trained? This section needs to be expanded considerably as described in the general comments.	A new section was added to the paper about the citizen observatory, which addresses these questions. Citizen data were not discarded but used to investigate their value in complementing hydrological modelling (Mazzoleni et al., 2017; 2018).
L 377: "The situation before the intervention" The word situation is rather vague, but more importantly you do not describe what exactly that intervention is (see general comments).	We have modified many headings to replace intervention with implementation of the CO on flood risk management.
L 379: Specify what you mean with numerical simulations. Do you mean the hydraulic model? Very little information is provided about the modelling that was done. The reference to section 2 could be more specific, do you mean section 2.2.1?	This has been modified to indicate numerical simulations from the hydraulic model, and the reader is then referred to the Supplementary Materials.
Table 8: What is P1, P2 and P3?	These have been replaced with the hazard classes of low, medium and high.
L 415: "implementation of the CO" -> not fully explained what that is, see general comments.	A new section was added to the paper about the citizen observatory.
Table 11: Please explain the difference to table 9 (not just the different values, but the different meaning).	The difference between the two tables is the amount of area flooded by risk class and flood hazard scenario before (Table 9) and after (Table 11) implementation of the CO. The captions have been modified to clarify this point and additional text has been added.
L 426-432: This paragraph is rather unclear, but this will likely improve once the general comments are addressed.	We removed this paragraph but provided more details on how the risk changes with and without the implementation of the CO in the methodology.
Table 12: caption "different" not "difference"; also I would not just include the residual damage, but also the original and then the difference. Or you could make a plot with both the original and residual costs.	This caption has been corrected. The original costs and the differences have been added.

Comments from Reviewer 1	Response to Comments
L 436: How did you calculate the cost for CO? What are the cost components? Is it sensor costs, maintenance costs or personnel costs? Over what time do these costs accumulate?	Table S2 has been added to the Supplementary Material, which is a breakdown of the costs of the CO.
I have found at least one reference in the text that is not in the list of references. Please double check.	All references were checked to ensure that they are in the reference list.