Response to RC2
We would like to thank the reviewer for their comments. These appear in bold typeface; our responses follow below each comment.

General comments
This publication by Ferri et al. discusses a cost-benefit analysis for citizen observatories based on a specific catchment in Italy. The content is relevant and will be a valuable addition to citizen science research. One of the current limitations of this paper is the lack of a broader context and the limited discussion. Questions that would be interesting to see addressed in the discussion section are: Why is it mostly the "social dimension of vulnerability" (L 354-355) that changes the calculations and not the additional data obtained through the CO? Is this additional data not helping to further improve the hydraulic model? What aspect of the "social dimension of vulnerability" do the authors contribute most of the reduction in costs to? This is slightly hinted at in the methods (L 357-362), but as one of the main messages in this paper this should be addressed more extensively in the discussion. A full discussion of the results and the broader context of the study would make the value of the publication clearer to the reader.

Response: We thank the reviewer for their positive comments regarding relevance and a valuable addition to citizen science research. We agree that the discussion could be improved to provide a broader context. For example, the additional data obtained through the CO are indirectly part of the improvements to social vulnerability rather than the hazard component that involves the hydrological-hydraulic model as one might be expecting. This is because the hazard modelling is done as a baseline while the data collected by the citizens/expert volunteers will be in near-real time so this will affect the early warning system, the involvement and hence awareness of citizens, and the updating of the hazard and risk information, which are all part of social vulnerability.

One of the aspects that contributes most to a reduction in costs is related to the **emergency services.** Although the people involved in the emergency services are the same, they are employed in a much more efficient way as a result of the technology developed within the CO, which allows for better management of the teams responding to the event and the assignment of tasks based on an operator's location. For example, the authorities, in response to real-time reports from citizens and based on reliable model forecasts, can assign critical tasks to rescue teams, monitor the movements of the teams and assign new tasks to teams once a job is finished based on the proximity job principle. Secondly, among the factors of social vulnerability, the involvement of citizens contributes to changing those behaviours that are the main causes of death and/or serious economic damage; an example would be actions such as trying to save your car during a flood event or trying to rescue your belongings from a flooded basement, which would be stopped or reduced as a result of the CO.

We will add this type of additional context to the discussion to address the limitation pointed out by the reviewer.

Specific comments:
Overall there are too many abbreviations (e.g. L 214, 258, table 6, L 369, L 379). I was not able to find the definition of the abbreviation "EWS" (table 6).

Response: Regarding L214, this is a blank line in our version of the paper unless you are referring to L215, which defines one of the variables in equation 2? Regarding L258, is the reviewer referring to FHR or Vp as an acronym? If it is FHR, this was defined further up in the text and used in equation 3. If Vp, then we

have now added this to the caption of Figure 4. In Table 6, we have changed the acronym EWS to read Early Warning System as this was not defined previously in the paper so thank you for pointing this out. On L369, CLC refers to Corine Land Cover, which we defined in L192/193. L379 defines variables that are then used in equations 5 and 6 and further in the text. However, we take the reviewer's point that there are many acronyms and variables defined and used throughout the paper. We would be happy to create a list of acronyms if this would help or remove some of the acronyms and replace them with the full text when they are not used very often in the paper.

L 55-59: Not all of the cited literature actually refers to a CO and the description of at least some of the stated studies is not accurate.

Response: We have checked the referenced literature and deleted those references that do not refer to COs (Etter et al., 2018; Mazzonleni et al., 2017; Butaert et al., 2014) but to other forms of citizen science. Moreover, we have attributed the statement regarding the link of COs with authorities and policy more specifically as follows:

"Specifically, Wehn et al. (2015) found that the characteristic links of COs to authorities and policy do not automatically translate into higher levels of participation in flood risk management, nor that communication between stakeholders improves; rather, changes towards fundamentally more involved citizen roles with higher impact in flood risk management can take years to evolve."

L97-100: How often do these observations get made and how many were collected in total? It would be very informative to include a photograph of such a "staff gauge with a QR code".

During the WeSenseIt project, more than two hundred people were recruited for practical activities and were trained to use the WeSenseIt technologies. The data collected took place during the evaluation exercises organized to test the technology and to collect feedback for further development and to make improvements. During these events, the response of the volunteers was enthusiastic as well as their participation in sending environmental reports and information. Examples of photographs are provided below, which we could add to the paper.

L 104: Did the volunteers operate the physical sensors? Or was this done by someone else?

Response: No, the physical sensors are operated by AAWA in collaboration with the Regional Department for Soil Protection, the Environmental Agency, the Civil Protection Agency and their related professionals.

L 197 / table 1: It is not clear to me which of these data inputs are derived from citizen scientists and which are implemented anyway. Please make this distinction clearer so that the added value is more obvious.

Response: None of the data in Table 1 (version posted online) are derived from citizen scientists. We have now added the input from citizen scientists to the Flood Vulnerability component as follows:

Table 1: Input data used to calculate risk.

Component of risk	**Data**	**Source**
Flood Hazard	Same as before	Same as before
Flood Exposure	Same as before	Same as before
Flood Vulnerability (Susceptibility)	Vegetation cover	Corine Land Cover 2006
	Soil type	Corine Land Cover 2006
	Water height from simple gauges equipped with QR codes, which are read by technicians and citizens as well as photographs and other flood-relevant information collected via an app	Collected by AAWA

These data are used indirectly in the calculation of social vulnerability, i.e., the Early Warning System (EWS) component through improvements in the reliability, lead time and information content of the EWS (Figure S2 in the Supplementary Information) as well as two components of Adaptive Capacity (Hazard and risk information updating and Citizen involvement - Figure S3 in the Supplementary Information).

L 425: It would be helpful to add a range to this value, so as to show the associated uncertainty.

Although we understand the point of the reviewer, we do not have a range around the expected average annual damage because it is already based on three probabilistic flood scenarios, so this damage value is an average across these scenarios. We recognize that there are multiple assumptions and uncertainties in this methodology, but we have not quantified them as such. Since the exercise is currently a hypothetical one but it provided sufficient evidence for funding of the CO for five years, going back and doing a full uncertainty analysis would not bring any further value for the operational running of the CO. Instead, once the CO is operational, it will then be more interesting for us to verify the results from the cost-benefit analysis.

L 430-432: Why do you think there is a difference, i.e. why is R3 and R4 reduced, but R1 and R2 increased? Add this to the discussion.

Response: This occurs because the total area affected by the flood hazard is the same before and after implementation of a CO. What changes is the distribution between risk classes, i.e., R3 and R4 are reduced, which means that the areas at risk in classes R1 and R2 will increase. We will add this to the discussion regarding the implications of this finding.

Also table 9 does not show any areas, just damage, so the reference here probably refers to table 8?

Response: Thank you for pointing out this error. We have corrected this but in the process of responding to reviewer 1, we have also combined tables, i.e., Table 8 and 10 are now combined (risk before and after the implementation of a CO in terms of area) and Table 9 and 11 (damage in euro amounts before and after the implementation of a CO). The text has also been updated accordingly.

L 469: You mention that this method can be transferred to different catchments. It would be interesting to read your thoughts on what type of catchments this would be suitable for, e.g. what catchment scale.

Once activated in the Brenta-Bacchiglione, the CO will also be extended to the other basins of the hydrographic district of the Eastern Alps, which are similar in size and hydrological characteristics. These are complex hydrographic basins with very variable regimes, from rapid response/torrential rainfall events of the alpine territories to the alluvial plain, which is composed of mountain and lowland river networks, artificial networks of reclamation and natural and/or artificial reservoirs. In general, we suggest applying this methodology to catchments larger than 100 km^2. For catchments of this size and greater, we would have basin compatible response times, and hence, it would make sense to use model forecasts.

Another point to note is that for the application of the CO methodology, it is necessary that the population residing in the basin can be easily reached through such an initiative, and that they are familiar with, and are able to access, the technology (i.e., via a tablet, PC, smartphone).

Technical corrections:
L 92: 7th (th in superscript)

Response: Thank you for spotting this error. We have now corrected this.