Generic:

This manuscript is a useful contribution to the field of hydrology and flood modelling. It also complements similar papers recently published which have used real crowdsourced/citizen science observations to support real hydrological applications. It is evident that the manuscript has been reworked considerably since the original submission and has addressed the reviewers' comments where necessary, which in turn has strengthened the quality of work. However, I still think it is confusing in places, particularly the methods section, and how the synthetic and real-time data were actually generated. I also think that a lot of assumptions are made (which are not always referenced or explained) without fully appreciating the complexity of engaging and involving real citizens. Furthermore, I do not think data quality is fully appreciated; for instance, physical and automatic sensors are subject to error throughout the data collection phase, and not just when level is converted to streamflow. Citizen-based observations are also subject to error in a number of places, particularly if relying on photograph submissions v's a quantitative value. The value of these particular scenarios are therefore limited. It is difficult to apply the findings to a broader scale and be used to influence operational activities. The manuscript reads well in places, but some sentences could do with being rephrased or reworked to improve fluency or clarity. The wider picture needs to be clear and reiterated throughout. The results and accompanying figures are well presented.

Specific (suggested additions in blue italics, text to delete in red italics):

Page 1, Title: consider changing 'a model study based in...' to 'a *modelling* study based in...'

Page 1, Title: consider either adding '*(Italy)*' to the end of the title so the reader knows where Bacchiglione is, or make it more relevant to a wider audience by removing the place catchment name altogether.

Page 1, Line 17: make this clearer. 'less accurate' because non-professionals / the public are collecting valuable hydrological datasets? It needs something to describe what crowdsourcing actually is within the abstract.

Page 1, Line 19/20: 'the extreme flood event *which* occurred in'

Page 1, Line 21: what do you mean by target point? Do you mean receptor or impact zone? It isn't a very common flood risk management term.

Page 1, Line 21: 'Ponte degli Angeli (Vicenza), *at the* outlet of the Bacchiglione catchment'

Page 1, Line 28/29: upstream sub-catchment scenarios are very catchment specific. It depends on whether you have a community upstream (less likely) or nearer the outlet (more common). This will affect your results.

Abstract: would benefit from documenting more/clearer results in the abstract.

Page 2, Line 1: find a better term for 'proper'. What does this even mean?

Page 2, Line 5: 'for example, to operate control ~~river~~ structures…'

Page 2, Line 6: 'Reliable *and* accurate streamflow simulation…'

Page 2, Line 7: 'inherently uncertain due to *the.* lack of reliable…'

Page 2, Lines 7-13: embed the list of points into the sentence better. E.g. use 'for example..'

Page 2, Lines 13/14: 'Data assimilation is a common *method for* updating model input.'

Page 2, Line 23: 'citizen-based data *(*Shanley et al…'

Page 2, Line 29: 'In both studies, the *observation* filtering process…'

Introduction: make it clearer what crowdsourcing actually is and perhaps link them to other similar terms e.g. citizen science ad VGI. Also consider that 'usefulness' doesn't just relate to flood forecasting and real time information. Crowdsourcing can also contribute to, or generate new, long-term datasets over time, and support other types of management activities.

Page 3, Line 4: avoid repeating the same word in the sentence ('mentioned')

Page 3, Line 7: how is your study real-time? It is not clear. There are reasons why real-time has not been focussed upon e.g. citizens submit their observations at a later date when they have phone signal, wifi or data to submit them.

Page 3, Lines 4-8: this text outlines the main research gap for your work (which is good) but are not clearly reflected in your abstract.

Page 3, Lines 10-13: '~~To that end, w~~We analyse a~~the~~ flood event *which* occurred in May 2013 in the Bacchiglione basin *(Italy)* derived from a distributed network of StPh, StSc and DySc sensors. Synthetic CS observations of water level are assimilated in a cascade of hydrological and hydraulic models since real CS measurement are not yet available *for this particular study site*.'

Page 3, Line 15: Useful to include a final sentence to say how your papers aim/objectives have a broader relevance.

Page 3, Line 18: Useful to include a link to the WeSenseIt project. Don't assume the reader knows what this is.

Page 3, Lines 23-25: The project set up a pilot – this is confusing. Makes it sound like citizens were actually involved. If they weren't, who was?

Page 3, Line 25: 'usefulness of assimilating CS *WL* observations ~~or WL~~ to improve the model performance and consequent*ly* flood prediction'.

Page 3, Line 26: 'Northern East' should be written as 'North Eastern'. But I would move this to Line 18, when the catchment is first introduced 'The Bacchiglione catchment (*North East* Italy)'.

Page 3, Line 27/28: 'river length of about 50km'. Use approximately instead of about.

Page 3, Line 28: change left side / right side to east? West?

Page 4, Line 1: Forecasted and measured precipitation time series – were these subject to quality assurance and control checks? Are they of a high quality?

Section 2.1: Do you need to refer to a location map within this section (i.e. Fig 1)?

Page 4, Line 6: 'Three types of sensors *used* to measure WL, static physical (StPh), static social (StSc) and'

Page 4, Line 9/10: Any quality control checks for the StPh traditional sensors? Sensors are still subject to error. Why assume? Why not prove this?

Page 4, Line 13/14: Is the mobile app used to submit photos, videos and/or quantitative values? Are date, time and location also submitted (i.e. metadata)? Any data quality checks anticipated/required? The app and use of QR codes is very specific and difficult to synthetically generate.

Page 4, Line 23: 'We assume a direct relation*ship*…'.

Page 4, Line 24: 'i.e. the probability of receiving a CS observation~~s~~.'

Page 4, Lines 16-18 / Page 9, Line 13: estimating velocity and runoff induces significant uncertainty and defeats the object of involving citizens in a cost-effective and simple way. Is it worth the effort if additional data is required or has to be derived? It is unlikely that rating curves would be available in reality. Some studies

are extracting velocities, levels and discharge from videos and photographs automatically using image analysis techniques.

Page 4, Lines 31/32: CS activities are not yet operational but this page describes these activities. This makes it confusing to follow. It is not clear how/if synthetic data is used.

Page 5, Table 1: How are the photos used? Who extracts the information? Social observations can come in a variety of formats, and is often one of the biggest challenges/barriers when involving citizens. How would this be managed in practice?

Page 5, Table 1: Why is StPh regarded as a CS method here? It is automatic and generates the data for you.

Page 5, Table 1: Do you have any references to add to the observational error column? Examples do exist in the literature and data quality is important.

Page 5, Line 9: 'from a wide range but limited number of' – this is not clear.

Page 5, Line 10/11: due to the limited number of participants – isn't that the point? Recruitment and low participation is a huge barrier.

Page 6, Line 8: 'In *the* case of the main river channel,'

Page 6, Figure 1: it would be useful to mark on the map where the urban area of Vicenza or 'target point' is.

Page 6, Line 14: 'Figure 1. Spatial distribution of the sub-~~-~~catchments, river reaches, *and* StPh and StSc sensors implemented in the catchment by AAWA'

Page 6, Line 18: 'relate to the model equation *here* as *a* detailed description is available in Ferri et al. (2012)...'

Page 6, Line 18/19: Precipitation time series – can/have the citizens observe this too? Many examples in the literature where they have.

Page 7, Line 3: '~~The t~~*T*emperature is used for the estimation'

Page 7, Line 7: Information on the quality/success of the calibrated model would be useful. Do you have any statistics to validate its performance?

Page 10, Line 4: I do not agree that rating curves are the only source of error/uncertainty. Especially when physical sensors often measure water level indirectly using temperature and pressure.

Page 10, Line 14: WL can be easily measured by citizens using a staff – this depends! Some studies have found that their ability to manually observe level using a staff can vary greatly. It can also depend on when it is installed, how turbulent the flow is etc. I feel as though any error associated with the citizens is completely bypassed here. It cannot be assumed that error is the same spatially, temporally and for each participant.

Page 10, Table 2: It would be useful to include a citation for the coefficients used in your study, within the table itself or within the table caption so it is clear when they have each come from.

Page 11, Lines 15-20: What NSE value do you regard as being 'good' or 'acceptable'?

Page 13, Line 17: why have you used 500m and 1000m? Why are they assumed? Citizens may travel or walk elsewhere.

Page 13, Line 29: 41% still seems very vague/generic in the context of data submission.

Page 15, Line 7: Batson et al 2002 seems an old reference to use for such an evolving topic which is heavily dictated and driven by technology.

Page 17, Line 8: Why have you used 80%?

Page 18, Line 6: 'and river reaches (hydraulic model) for *a* 1-hour lead time.'

Page 18, Figure 4: would be useful to include 'NSE' on or next to the colour ramp key. And repeat for all later figures.

Page 27, Line 11: 'so for the assimilation of CS observations it is *also* important to consider ~~also~~ this'

Page 28, Line 10: 'This section aims *to summarise* ~~at summarizing~~ the main findings of our study and…'

Page 28, Discussion: there is scope to relate your findings to the literature in more detail, including those which have used real crowdsourced observations.

Page 29, Line 23: 'awareness of flood risk decreases over time' – do you have a reference to back this up?

Page 29, line 28: 'Gharesifard and Wehn (2016) *are* ~~and~~ Rutten et al. (2017) and being studied in detail in the H2020 GroundTruth..'

Page 29, Line 32: 'This study demonstrates that high *performance models* ~~value of model performance~~ can still be achieved even…'

Page 30, Lines 7-9: This text is not reflected in the abstract, despite its importance.

Page 30, Line 13: Why discuss experiment 2 here and not experiment 1?

Page 30, Discussion: what do your results/conclusions mean for the wider picture? Ensure readers can relate to your study.