Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/nhess-2021-169-RC1, 2021

Comment on nhess-2021-169

Anonymous Referee #1

Referee comment on "Geo-historical database of flood impacts in Alpine catchments (HIFAVa database, Arve River, France, 1850–2015)" by Eva Boisson et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2021-169-RC1, 2021

This paper presents the establishment of a flood impact data set in the French Arve Valley. While the data collection and preparation is commendable, the analysis itself and the presentation quality is rather basic. In essence, the manuscript presents an interesting data set and some simple descriptive statistics of this data set. If this is intended as a data paper, and the manuscript serves as a reference for describing a data set and its collection, this would be fine. Otherwise, I would argue that the exploratory analysis of the data set could be improved in order to foster scientific novelty.

My comments and suggestions for improvement are detailed below:

Abstract

I suggest to streamline the beginning of the abstract, which is a little bit clumsy.

Some examples:

- l. 24: suggest to specify "it"
- l. 25: the "Alps that warm at a rate twice as high in the Northern Hemisphere" - although this will probably be understood correctly by informed readers, the statement itself is somewhat unclear

Introduction

- l.118: I suggest to be more precise as far as the "first database documenting a mountainous catchment" is concerned. There are definitely databases on historic floodings and debris flows in other countries in the European Alps, including Switzerland, Austria and Italy. These do comprise natural hazard events on a catchment level, and some have been linked to exposure and mitigation (e.g. https://doi.org/10.1016/j.crm.2021.100294 for torrential flooding; https://doi.org/10.1016/j.gloenvcha.2020.102149 for avalanches), therefore effectively covering the "interactions between social and natural dynamics engendering flood impacts".
- The term "impact" is a core concept throughout the manuscript. I think a clear definition on what an "impact" actually comprises would be helpful (either in the introduction or in section 3). This is somewhat hidden in section 3.3, I suggest to state this more clearly earlier on.

Study area

- Figure 1: Please rework the elevation legend. This is a continuous scale and should be presented as such. If a discrete scale is used, intervals need to be reported and not scalar values. Which color does an elevation of say 1000 m a.s.l. correspond to? Is it green or yellow, or something in between? In addition, I would advise to use prettier breaks, and not 401 - 1401 - 2902 - 4810, which seems somewhat arbitrary. The color of the Arve river in the legend does not correspond to the color in the map.

Materials and methods

- l. 254: ... can not be estimated by ...
- l. 255/256: "The most recent sources are often highly informative, allowing impacts to be more precisely located"
- l. 285ff: This statement belongs to the "Outlook" section.

Results and discussion

- Figure 2: The caption could be more informative. I assume that ticks indicate mentions, and the shaded area desplays cumulative mentions, but this is somewhat speculative.
- Figure 3: I suggest to rework this plot completely. First of all, using x-axis ticks would be helpful here. The dense number of x-axis labes does not help, since single bars are difficult to assign to specific years. Also, the second y-axis is confusing for two reasons. First, a transformation between the two y-axis of f(x) = 0.266666 * x has been used;

consequently only 8 and 16 share a common y-grid line with the primary axis. All other labels float around somehow. Secondly, I do not understand why moving averages (especially for the impacts) were put on a secondary axis. Plotting data and a smoothed version of the very same data on different scales in not intuitive.

- Figure 4: You could use some alpha/transparency for plotting the location of the impacts. Also, the very colorful background makes it somewhat difficult to discern the different colors. Personally, I find simple bar plots to be easier understandable than bar plots in polar coordinates (i.e. pie charts). Also, point types to not match (circles on the map, ellipses in the legend).
- l.378: The authors write that "The increase in the number of impacts starting in the 1920's and well-marked from the 1960's can be explained by multiple factors such as indirect source effect, increasing flood activity and/or increasing exposure of goods and people." It is hypothesized that this is attributable to increased exposure or the evolution of data sources, but there is no proof for these statements. One core aspect I am somewhat missing in the discussion here is the completeness of the database. It has been shown that underreporting of events is likely in the time period up to the end of WWII. Has this been taken into account, or is the dataset simply assumed to be complete?
- l.381: "Therefore, the increase in impacts cannot be explained by changes in flood occurrence, at least prior to 1990." This is an interesting finding, that of course needs to be discussed. The authors provide some explanations in the following paragraphs. However, I find these paragraphs a little bit difficult to follow. I think this section needs to be reworked with a focus on clarity, e.g. by providing a table detailing the sources per period for easier comparison, or providing some sort of visual emphasis on the main lines of thought here. Ultimately, upon reading the paper I am not sure where these observed trends in impacts do come from?
- l.400: "We can assume that, floods are more likely to be reported in newspapers as when they happen in a location known by the reader." Could the authors elaborate this? I would assume that this does not affect events of a certain magnitude?
- l.403: illustrated by Figure 5.
- Figure 5: See comments on Figure 4. The colors of the sources are even more difficult to spot here, due to the colorful background. I suggest to use a more neutral background, the important information is contained in the points, not in the elevation.
- l.435: "We can see that the trends are significantly the same." Apart from the fact that trends can only be significanlty different, but not significantly the same (from a statistical point of view), I would like to point out here that - naturally - impacts and events show a higher correlation in the Figure from Appendix 4 than in Figure 3.
- l.440: Changes in exposure and vulnerability are only briefly (basically 1 paragraph) discussed in a qualitative way. I assume that data for providing a more detailed assessment of this line of thought is not available?
- Figure 6: top right: again, secondary y-axis is somewhat difficult to read, as 0:120000 is mapped to 0:7. I am not really convinced by these plots with secondary y-axis, especially since there is no natural relationship between these two data sources. The right axis can be plotted from 0:6 (spreading the range from a visual point of view) or from 0:12 (squishing the impacts in relation to the population count), but I am not sure which would be the more "correct" one. At least, I suggest to match the color of the y-axis to the line color.
- l.475: "...height are only mentioned in rare cases."
- l.481: "Impacts on industrial facilities"
- Figure 7: Suggest to try a mosaic plot for visualizing these data.
- l.507: The categories used for the analysis
- l.509: Please clarify what "evolution of the assigned words" means.
- Figure 8: Histograms usually do not have whitespace between the bars, as they represent continuous variables. Also, I suggest to work on the colors, as impacts have a similar color as victims, and events have a similar color as protection infrastructure.
- l.515: hinders: do you mean hides? covers? conceals?

- l.517: what is meant by "(16 out of 28)"? I assume an increase from 16 to 28, and not the proportion 16/28 of something? Please clarify.
- l.515-520: I do not fully understand the point the authors try to make here. The "slight increase" (which could also be considered as a not-so-slight increase of 75%) is hardly visible in the plot, because the share of the "victims" category is quite small altogether? This is a rather trivial observation, and in fact just a matter of data presentation.

Conclusion

- Again, the authors take up the hypothesis that the observed increase in impacts could be explained by exposure and evolution of sources. While I tend to agree in principle, I would like to emphasize that this is not really shown in the paper, but remains a hypothesis. Other aspects such as mitigation measures (either technical ones or soft measures such as awareness raising) are not considered. The effects of data completeness (i.e. underreporting in the earier years of the time period), effects of exposure and land-use are not investigated in detail.