Referee 2

This paper presents a thorough review of the ORRION database, it details the content, spatial and temporal extent of the database and provides some preliminary analysis of the results. The paper is of particular interest to those interested in the Alsatian region, but more could be done to highlight why this is of interest beyond this particular case-study. I have added a broader range of references to the attachment which the authors may find of value, many more are available beyond these which would strengthen some of the arguments you make. There are some interesting points in the paper, but these need to be clearly teased out in the conclusion, as it is a long paper.

Authors' response: Thank you very much for your interest in our work, your positive and constructive feedback and your detailed review. All suggested references and minor corrections proposed in the attached document will be taken into account in the reworked version of the paper. Also, the conclusion will be slightly reworked to better highlight main results and conclusion of our work. We agree that this is necessary given that the paper is long due to the very large amount of information it conveys regarding both the investigation method and the data summed up. Especially we will make a bit clearer how the work conducted already contributed to very concretely improve flood risk management in the targeted area. By the way, below are our responses to all the comments and questions raised.

I was surprised that the paper did not reference any previous work using the historical descriptors as either a proxy flood level indicator and in indices production, I have added suggestions from the literature, this may just be an oversight in writing the paper up, with this undertaken during the development phase. It would be good to know how you accounted for recorder bias through time, e.g. a potentially more interested individual at a site, and changing literacy rates.

Authors' response: Thank you very much again for all the suggested references. They will all be considered in the revised manuscript. Regarding the question of the source bias, we want to stress that that our paper aims at introducing and describing the ORRION database. ORRION's goal is to share raw flood information only, at least in its current state, avoiding subjective choices as much as possible, with the ultimate objective to provide in a transparent manner all the information required to conduct a geo-historical analysis of the event chronologies. Such an analysis allows to highlight the factors that explain the spatio-temporal distribution of flood events, namely anthropogenic factors related to sources, vulnerability and infrastructures, but also natural factors related, e.g., to climate. This is discussed in section 5.1 of the paper, that highlights the major patterns of change that are visible at the scale of the entire database, and/or more specifically for certain rivers and locations. For instance we clearly say that the shape of the records (number, damage types) reflects, among other effects, the availability of sources as function of time. Hence, we do not correct the record but we tried to make the record as exhaustive as possible, and we reports all sources related to events within the database. These remain available to anyone interested. Indeed, ORRION data is based on an extensive archival campaign (Transrisk project) further complemented by participative data. This was made in order to maximize the amount of information summed up regarding past events. In addition, the evolution of sources with time at regional scale has been already studied by Himmelsbach et al (2014) and, regarding specific bias, by Martin et al. (2015). This all, in addition to the complete availability of sources within the database, should enable the source bias to be corrected in the future by anyone interested in using the information for, e.g., improving risk mitigation or studying the influence of climate on flood occurrences.

This point, already present in the current version of the paper will be highlighted even more strongly in the reworked version of the paper.

I am very surprised that no attempt at flood reconstruction is undertaken, either in a simple e.g. level to discharge estimation, or in a more advanced manner (Herget & Meurs, 2010). Surely the documentary evidence would permit such in some instances, particularly if sufficient information to map flood extents. If this is planned state that, as this is of value to flood risk managers. It would be good to see some comment on the influence of the engineering activities over several decades on some of these systems in more detail – can they account for the changes? Or guide reader to other work that has undertaken this.

Authors' response: Indeed, the information gathered makes it possible to reconstruct peak discharges (more or less precisely). Reconstruction methods are now available in the flood literature and are becoming more and more popular. Annette Boessmeier developed and used some of these in her PhD to reconstruct peak discharges for the Kinzig river belonging to the German part of the TRansrisk project. Also, ORRION's information already allowed to better characterize the mapping of the historical floods of 1852 and 1860 for the Ill in Mulhouse, and to validate the corresponding discharge modelling that has been realized within the framework of the European flood directive. Finally, the study of old water levels already revealed that the computation of the centennial flood of the Zorn river was affected by several errors (Martin et al, 2018). Hence, we are well aware of the potential of discharge reconstructions and of the usefulness of our data to achieve this goal. Simply, doing this for all our rivers is beyond the scope of the work presented here. Annette Boessmeier did it for one river only, and this took her most of her PhD… Here, the objective is more modest: to introduce what we already have (a lot of information and chronologies of events, typologies, etc.) which leads to an article that is already long (cf. referee 2 review.). However, of course it would be very useful to do the discharge reconstruction all over the study area and period, and this is clearly an objective for future developments. The reworked version of the paper will precise this and insist on the promising perspectives offered by discharge reconstructions on the basis of the available information.
Regarding the question of the impact of engineering activities including construction of mitigation measures and more widely of land use on the flood records, as for the sources, this is already in the discussion section of the paper where we provide quick explanations for some changes we observe. This applies, e.g., to flood chronologies with/without overflow for the Rhine and the Bruche rivers. Clearly, more could be done at the fine scale of specific municipalities or rivers and such analyses were already done in some cases. For example, Martin et al (2017) demonstrated the impact of the construction of the Ill canal at Mulhouse and in Strasbourg on the flood record. At a more regional scale, it could be shown that the spatial distribution of the consequences of the 1910 flood in the Rhine plain can only be understood if one considers the uneven development of river correction works at that time. However, conducting such analyses is extremely time consuming and could not, so far be achieved at the scale of our entire database. This point, already present in the current version of the paper will be highlighted even more strongly in the reworked version of the paper, especially in the method (introduction of ORRION) and discussion (interpretation of events chronologies) sections.

I felt the discussion would benefit from some comparison of the flood rich years that emerge potentially being discussed in more detail and comparison with other databases on long flood series from across Europe - no more that that is possible at present until underlying trends within the data resulting from social factors are addressed – the lack of this limits the utility of the results to a database compilation.
Authors' response: This is a good suggestion, thanks. In the paper, we already briefly compare the smooth temporal trend in our records with other examples of the literature (e.g. Wilhelm, 2012). However, comparing specific flood years / events is another appealing option. We therefore compared the most severe events of our record (class III and IV) to the

extreme flood events identified by Glaser et al. (2010) for 12 major rivers of central Europe including the Rhine. Specifically, Glaser et al. (2010) identifies 13 major flood events between 1500 and 1900. Among these, two correspond to high (class III) or extreme (class IV) events of our record: January-February 1862 (class IV) and February 1876 (class III). Another event of Glaser et al.'s inventory is for us a class I event (march 1784) whereas 4 other events of Glaser et al.'s inventory correspond to events in our record, for which the severity class could not be determined using our specific scale (this is the case for 35% of the studied ORRION events in total). Considering this latter specific difficulty, the comparison appears therefore as rather probative, with around half of the years of Glaser's inventory corresponding to ORRION events. This is all the more true when considered that, whereas the work of Glaser et al. focuses on major rivers of the length and watershed side corresponding to the ones of the Rhine, ORRION events correspond to a very wide range of flood types in terms of concerned rivers and spatial extent. Especially, a significant part of ORRION events concerned only one or few "small" Alsatian tributaries flowing down from the Vosges mountains, and that occurred after a very localized rainfall event such as a thunderstorm. It is therefore not surprising at all that the agreement between Glaser et al.'s and ORRION records is far from perfect. Finally, for the 20th century, correspondence between well-known flood events and our record can also be found. January 1910 was the most intense flood event of the Seine river in Paris over the last centuries. It is also a class IV event in our record. January 1936 that was the time of an extremely widespread and intense flood episode over the north Atlantic (Ballesteros Canovas et al., 2018) is also in our record, but corresponds to an event for which the severity class could not be determined. All in all, even if, of course, much more could be done in terms of comparison between ORRION and other records, this quick comparison already further demonstrates that, even if the event chronologies that can be derived from ORRION database are clearly affected by human factors related to sources, vulnerability etc., they also include information related to the physics of the flood process and its underlying hydro-meteorological causes.
This discussion section of the reworked article will include this quick comparison as an outlook, among others, offered by the data summed up in ORRION.

Further discussion on quality control of public database contributions is needed, a statement on how this was managed.
Thank you for this question. The quality checking is indeed a crucial point for the reliability of the database, and also a huge part of the work of the database management. Basically the team of database administrators do this by verifying possible redundancy with pre-existing information, directly interact with the public information providers, check themselves the provided sources (e.g. pictures, newspaper articles, etc.), and, if necessary, go themselves on the field for further verifications or do themselves further archival searches. In practice, the likelihood of many testimonies is to be investigated in details, and also retrieving the date and location of the pictures provided is often a huge work. This information will be added in the revised manuscript.

p.5/6 I would benefit from some indication of size of river discharges being considered, I don't recall seeing anything of this nature – consider adding.
Herget, J., & Meurs, H. (2010). Reconstructing peak discharges for historic flood levels in the city of Cologne, Germany. Global and Planetary Change, 70(1–4), 108–116.
You are right considering the interest of a brief river description. Basically, the Rhine river is a well-known very large river, the Ill river is one of its important tributaries whose basin encompasses most of the studied area, and the 11 other rivers considered are smaller rivers, most of them being tributaries of the Ill river. The following table sums up the characteristics of these 13 rivers in terms of length and 10 and 50 year return period peak discharge as considered in flood prevention plans. This information will be added in the revised manuscript.
By the way, the suggested reference (thanks) will be added in the revised manuscript

River (municipality)	Length (km)	Q_{10} (m^3/s)	Q_{50} (m^3/s)
Ill (Osthouse)	216	200	260
Largue (Spechbach)	51	61	82
Doller Burnhaupt)	47	62	84
Thur (Bollwiller)	53	91	120
Lauch (Guebwiller)	47	22	30
Fecht (Ostheim)	49	91	120
Giessen (Sélestat)	34	80	110
Lièpvrette (Liepvre)	25	34	48
Bruche (Oberschaeffolsheim)	77	140	190
Mossig (Soultz)	33	15	20
Zorn (Waltenheim)	97	79	110
Moder (Schweighouse)	82	68	93
Rhin (Lauterbourg)	1233	3900	4600

Please also note the supplement to this comment: https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2018-210/nhess-2018-210-RC2-supplement.pdf

Authors' response: Thank you very much again for all these suggestions (references, typos corrections etc.) and editing comments. They will all be considered in the revised paper.