Cover letter for the revised version of our article: "*Improving the understanding of flood risk in the Alsatian region by knowledge capitalization: the ORRION participative observatory*", submitted to Natural Hazards and Earth System Sciences (R2)

Dear Natural Hazards and Earth System Sciences editorial board, dear Dr. Uwe Ulbrich,

Thank you very much for your kind consideration of your work, and for the positive feedback about our revision. We provide in what follows an answer to the editorial report (which sums up the comments made by the referee) in which we detail that the requested changes have been made.

We hope that our new version that we are glad to submit to your kind consideration will be found adequate for publication in Natural Hazards and Earth System Sciences.

With best regards,

Florie Giacona, for the authors.

Editorial report

Dear Dr Giacona and Co-authors,
as you saw, the reviewer suggests the following alterations before the manuscript is published:
1. conditional means/probabilities instead of biserial point correlation,
2. keep the full categorical variables in the data set published instead of their representation by a set of binaries (which you might do anyway),
3. rescale the moving averages to the same units as the observations.

I am sure that your paper would benefit from obaying all 3 points, but given your focus on making the dataset available rather than performing a thorough statistical analysis, I leave the decision to you to folllow these recommendations, except that I demand one example of a contigency table (or respective text) as suggested by the reviewer in point 1 ahead of your correlation analysis (p16, line 9). This is, demonstrating the conditional frequencies for the combination of flood events in the rivers mentioned in the subsequent text lines. From the contigency, you can then shift to your correlation approach.

Author's response

Thank you again for your interest in our work and for your constructive feedback. Following your recommendation, we have added one contingency table (now Table 5 of the paper) to the paper. It is placed at the first time in the paper where linkages between a pair of variables is investigated. As this pair is one continuous variable and one presence/absence factor, we have added the one way ANOVA decomposition, its significance testing and the further test of the significance of the difference between both conditional means. ON the basis of the results obtained and their probative comparison to what is lead by the correlation analysis, we then say that we focus on it in the rest of the paper.

The approach is introduced this way in the methods section (p. 13, l 11-13):
"After checking on specific cases that standard variance decompositions / conditional mean evaluations led similar conclusions (Table 5 and Sect. 4.1), we evaluated association between all of our variable pairs with the classic Pearson cross product-moment coefficient." (p. 13, l 11-13)

And corresponding results in the first paragraph of the result section (p14, l22 – p15, l7):
The events mentioned by a large number of sources are floods that caused damage (especially affecting humans) and that had a large spatial extension. Specifically, source numbers per event are clearly different as function of the presence/absence of human damage (at least one people killed). Indeed, the one-way ANOVA decomposition of the source number per event variable is highly significant (p-value of the Fisher test $< 10^{-4}$), rebutting the assumption of a common population. Similarly, the conditional means of both populations (23.1 sources per event with at least one people killed versus 4 sources per event with no casualty, Table 5) are clearly distinct (p-value of the Tukey test $< 10^{-4}$). Logically, the Pearson cross-product coefficient between the variables "source numbers" and "presence of human damage" is +0.26, non-zero at the 0.05% significance level. According to these converging results, in what follows, association between variables is highlighted on the basis of such significantly positive/negative non-zero coefficient values only. For instance, all damage types are positively associated with the "source number" variable. There is also a very strong association between the source number variable and the extreme-sized events (class IV) and large-sized events

(class III) variables (Pearson coefficient: +0.56 and +0.29, respectively). Finally, it should be noted that the events mentioned by a large number of sources affected both the Rhine and its tributaries in particular (Pearson coefficient with RT type: +0.27).

	Number of events	Fraction of events	Mean source number per event
With human damage	31	10.5%	23.1
Without human damage	263	89.5%	4
Total	294	100%	6.9

Table 5: Source number related to each event as function of the presence/absence of human damage (at least one people killed). Only the events for which both the source number and the human damage presence/absence information are known are considered in the counts and for the evaluation of conditional means.

.

Regarding the full categorical data, yes, obviously they remain available, and information is indeed stored in the database that way. The recoding was just a convenience post-processing for this specific study which does not affect the original data.

Regarding the representation of the moving averages in the same units than the observation, we tested it and it leads to barely unreadable captions due to the high number of years with zero observations. We therefore keep our choice, given than the way we scale the moving averages is clearly detailed in text and remembered in the figure captions (“rescaled moving average”).