

## 1 **Reply to Anonymous Referee #2**

2 We thank the referee for his/her valuable comments. In the following we include our replies  
3 to the referee's comments directly at the respective points in the text. Referee comments are  
4 kept in italics and our replies are kept in normal font style. For better readability, we group  
5 the referee comments into topics, as some comments occur repeatedly in the referee's  
6 discussion.

7

8 *GENERAL COMMENT: This article deals with the usefulness of the "grey literature" in the*  
9 *study and analysis of natural hazards, especially in the case of floods in Germany. The*  
10 *authors analyse a series of documents and data extracted from various sources of information*  
11 *not covered in scientific journals. To do that, authors use an evaluation procedure based on*  
12 *the score of different parameters. The methodology is proven by a nonparametric statistical*  
13 *test to see the plausibility of the findings from two experts responsible to score a sub-sample*  
14 *of the studied documents. The work is innovative and addresses a current issue, because there*  
15 *are many studies published outside SCI journals routes, made by excellent professionals who*  
16 *for one reason or another (i.e. lack of means, tradition or linguistic problems) have not*  
17 *published their work in international journals. For this reason, I deem the publication of this*  
18 *work and opening up a wider debate about the usefulness of this type of information in*  
19 *scientific works. It should assume that the use of grey literature is subject to the control of the*  
20 *author and, therefore, they are responsible for checking the veracity of their references. I*  
21 *agree with the use of this type of information (after his selection) in scientific scopes.*

### 22 **1.1 SPRC concept**

23 *An example here is the called methodology: sources-pathway-receptors and consequences:*

24 *What's different between this method and the classical method of analysis of events (and*  
25 *widely used), which always studied the trigger, the process and their consequences?*

26 *Page 151, line 20-21 what's new in this concept?*

27 We use the SPRC framework for defining the dimensions under the contextual quality  
28 category. The SPRC concept is not new, i.e. it has been proposed by Samuels and Gouldby in  
29 2009 and used in various studies/contexts. For our purpose the framework provides the best

1 way of separating the hazard and vulnerability components, i.e. by identifying the sources of  
2 an event, its pathways and consequently the affected elements at risk and resulting  
3 consequences. In that it provides a standard for the multiple versions of how to separate or  
4 define trigger, process and consequence and largely helps to categorize event information  
5 under a unifying concept. The concept is easy to adapt to any natural hazard as it describes  
6 features common to natural hazards. E.g. for earthquakes it is possible to define the sources as  
7 the respective fault or slips that cause an earthquake and the pathways describe the attenuation  
8 and amplification of shock waves as they reach a built environment.

## 9 **1.2 Kappa test**

10 *Methodologically, I think the statistical test used may be acceptable although very few experts*  
11 *were selected for reviewing the sub-sample of documents. Do you believe that more experts*  
12 *could change the score-results?*

13 *Page 161. Line 2: 2 scientific expert is a low number for statistical analysis. Perhaps the*  
14 *results could be biased. If the authors can not improve this point, at least should address this*  
15 *issue in the discussion methodological*

16 The concordance check in the form of the kappa-test is an important part of the methodology  
17 we are proposing. It is commonly used in systematic reviews, i.e. in meta-analysis, both in  
18 clinical studies as well as environmental studies. We adopted the method for our purposes  
19 from these well developed study designs and respective guidelines (Higgins and Green,  
20 2011;Centre for Evidence-Based Conservation, 2010). Within these studies often only a  
21 limited number of experts (ranging between one and three) are used to test a particular  
22 selection scheme for its objectivity. Therefore, the reliability of the test rather results from the  
23 size of the sample given to the peers (assuming their expert status leads to a qualified  
24 judgement).

25 In our case we chose two independent and experienced scientists from the field of flood risk  
26 analysis. We deliberately did not use experts from the field of water authorities or other  
27 experts that largely publish their work as grey literature. In this way we are fairly confident  
28 that the critical scientific (and therefore event report users and not producers) perspective  
29 gives us an unbiased result. Certainly, increasing the number of experts would provide us with  
30 a more robust result, however, the peer review process is quite time consuming, so for the  
31 moment extending the expert panel is not really an option.

1 Our sample size is comparatively small, i.e. in total 10 studies that were drawn randomly and  
2 that constitute a good representation of the type of documents contained in the dataset. In our  
3 discussion of the results we elude to the rather small sample size (pgs. 161-162). Given the  
4 high kappa coefficients within the small sample we expect that increasing the sample size will  
5 not significantly alter the overall kappa, however, it would improve the statistical robustness.  
6 We will stress this aspect more prominently in the discussion.

### 7 **1.3 Minimum number of reports needed**

8 *On the other hand, one expects to see an analysis of what is the minimum number of*  
9 *documents that must be scanned for a significant improvement on the information. Based on*  
10 *your work, Is it possible to do this analysis? (I think that it is depending of different things,*  
11 *like the number of existing documents, the natural processes, the origin, etc: : :).*

12 The number of reports available, the geographical space covered, their amount of information,  
13 and their quality can vary largely for each event. Therefore, there is no unambiguous answer  
14 to the question. I.e., one high quality national report may provide enough information to  
15 understand an event in its extent and SPRC. In contrast, a number of qualitatively low or  
16 rather regional/local reports may still result in major gaps for event analysis. From the events  
17 analysed we cannot identify a general rule.

### 18 **1.4 Transfer of method to other processes**

19 *Finally, authors are really focusing on floods, do authors believe that the methodology is*  
20 *directly extensible to other processes?, or it is depending on the type of process (I mean, the*  
21 *evolution of the state of knowledge of the various natural processes could affect the*  
22 *methodology).*

23 There are two versions on how to read the questions (we are not entirely sure which is meant  
24 by the referee; therefore we will discuss both versions).

25 1) Is the methodology transferable to other natural hazards (NH)? – The quality assessment  
26 framework (QAF) for event documentations is designed to be universally applicable to any  
27 NH. The 4 quality categories are therefore defined hazard independently (s. pg. 151, from line  
28 8). The contextual quality category is also universal in its separation of the natural hazard  
29 components into the 3 dimensions sources, pathways, and receptors/consequences. Flood  
30 specific aspects are then introduced in the definition of the dimensions and their scores. These

1 definitions have to be adapted if documentation on NH other than floods is to be analysed for  
2 their quality. We believe this is a straight forward task and that the QAF provides a generic  
3 framework including the advantage of comparing the quality of event documentation across  
4 different perils.

5 2) Will an improved knowledge on any (flood) risk process influence the methodology? –  
6 This question relates to the contextual quality category and its definition of the dimensions  
7 and more so of the level of detail and type of processes considered. Again, the definitions of  
8 the three dimensions (sources, pathways, and receptors/consequences) are generic for floods,  
9 i.e. they provide a broad classification of the main governing processes and resulting  
10 consequences and therefore provide the guideline of the points that need to be considered in  
11 event analysis (and that therefore should be contained in an event documentation). For the  
12 example of the dimension ‘Source’ these are established categories like atmospheric  
13 processes, catchment state and runoff processes. Improvements in knowledge are to be  
14 expected in the degree to which these processes are understood and, hopefully, physically or  
15 statistically formalized. Our definition of the scores eludes to extract exactly that information  
16 from the reports that will assist in improving the understanding of the governing processes.

17 Will affect the result rather than methodology

18 Event analyses are an important step towards an improved understanding of the underlying  
19 processes (both physical and social) of natural hazards. Besides synthesising all available  
20 knowledge on floods, hopefully our framework will also contribute to a more systematic  
21 survey of key aspects of future events.

## 22 **1.5 Specific Comments**

23 *Page 147, line 14. May be palaeohydrologist methods can be here also cited.*

24 We refer to the most important works in the field of historic hydrology that consider  
25 documentary evidence. We will highlight more clearly that the study by (Brázdil et al., 2006)  
26 provides a most concise review of the works presented on historic hydrology for flood risk  
27 assessment in Europe and can be used as reference for finding further studies. Concerning  
28 palaeo-hydrology, to our understanding this discipline uses physical, chemical and biological  
29 proxies that can be related to flood events (i.e. using varved sediments of lakes to identify  
30 years/seasons of flood occurrences). Therefore, we considered to not include a review of these

1 studies in our introduction as we are interested in the human documentary evidences and  
2 methods of source criticism.

3 *Page 150, line 3-5. Why? They are not interesting?*

4 The analysis presented in this work are based on the document dataset by (Uhlemann et al.,  
5 2013) that is the results of a systematic search for flood relevant literature on trans-basin  
6 floods in Germany for the period 1952 – 2002. The methodology and rational for deriving this  
7 dataset is described in detail in (Uhlemann et al., 2013). The rational for limiting the search to  
8 particular types of documents is based on consistency criteria. These are ‘Scale and spatial  
9 consistency’ and ‘Accessibility consistency’. The scale consistency has implications on the  
10 choice of search tools, the languages for conducting the search and the types of references. In  
11 particular, the search can only be consistent (in Germany) at the level of white literature and  
12 publicly accessible grey literature of the higher governmental administration and national or  
13 international institutions. In terms of accessibility, to be consistent with the daily scientific  
14 search routines the search tools for the collection needed to be readily available also to any  
15 other researcher and at adequate expense. Therefore internet and media sources other than of  
16 public authorities’ or scientific institutions’ origin as well as narratives were not included in  
17 the compilation of the document data base.

18 *Page 150, line 25: user is needs, instead user’s needs*

19 We rephrase this part so it reads: “... from the perspective of the needs of the data user.”

20 *Page 153, line 15-28: It looks like an introduction paragraphs. Please consider to move it to  
21 this section*

22 We will remove this part from the methods sections and homogenize it with the introduction  
23 where the last paragraph on pg. 148 already outlines the argument.

24

## 25 **1.6 Length and structure of the paper**

26 - *You are using many times i.e. and parentheses, even though the information is also in the  
27 appendix section; you also include it in the text. I think that you can delete information  
28 repeated.*

1 - However, I think the article is too long. This sometimes makes it difficult to follow. I  
2 encourage the authors to do an exercise in synthesis and eliminating repeated paragraphs.

3 - In my opinion, the summary section is not needed if authors make a great effort to reduce  
4 the size of the manuscript, because they will get a more readable work.

5 - My decision: Accepted after minor revision. Authors should reduce the length of the  
6 manuscript and delete repeated paragraphs. One way to do it would be to remove information  
7 from the manuscript, because there is a lot of information in the text should not appear again  
8 in the text (i.e. a detailed explanation of each parameter analysed).

9 We will reduce the length of the paper. In particular, we will consider the referee's advice of  
10 reducing redundancies as they appear between the text and the appendix and reduce the  
11 explanatory sections in the methodology section to a minimum. We will keep the  
12 summary/conclusions sections, however, the section will be modified according to the first  
13 referee's suggestions (i.e. provide a more generic rather than flood specific discussion).

14

## 15 **References**

16 Brázdil, R., Kundzewicz, Z. W., and Benito, G.: Historical hydrology for studying flood risk  
17 in Europe / L'hydrologie historique pour une meilleure connaissance du risque inondation en  
18 Europe, Hydrological Sciences Journal, 51, 739-764, 2006.

19 Centre for Evidence-Based Conservation: Guidelines for Systematic Review in  
20 Environmental Management, Environmental Evidence: [www.environmentalevidence.org](http://www.environmentalevidence.org),  
21 2010.

22 Higgins, J. P. T., and Green, S.: Cochrane Handbook for Systematic Reviews of  
23 Interventions. Version 5.1.0 [updated March 2011], The Cochrane Collaboration  
24 (<http://www.cochrane-handbook.org/>), 2011.

25 Uhlemann, S., Bertelmann, R., and Merz, B.: Data expansion: the potential of grey literature  
26 for understanding floods, Hydrology and Earth System Sciences, 17, 895-911, 10.5194/hess-  
27 17-895-2013 2013.

28

29