# Response to Anonymous Referee #2 RC2: nhess-2018-407-RC1, 2019

Author: Jonas Laudan[1]
Co-authors: Gert Zöller[2], Annegret H. Thieken[1]

[1]University of Potsdam, Institute of Environmental Science and Geography, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany
[2]University of Potsdam, Institute of Mathematics, Karl-Liebknecht-Strasse 24–25, 14476 Potsdam, Germany

We thank the reviewer for the constructive comments. We will extend our descriptions analysis and discussion as suggested by the reviewer to improve the paper quality.

## Comments of the reviewer

**Reviewer quote, paragraph 1:**

A general limitation of the study seems to be that the people who experienced river floods experienced them multiple times in the last 10 years. While the people who experienced flash floods seem to have fewer past experiences. Is there a possibility that this frequency of past experiences may be a stronger signal than the flood type? Is there a way to correct your data for the number of flood experiences people have had?

**Answer to paragraph 1:**

We agree that this is a general limitation of the study since the previous experience of flash floods is very low among the surveyed residents. Yet, the residents who have been affected by river floods experienced several major flood events in recent years (2002, 2006/11, 2013) and therefore show a higher experience in total. However, not all regions that were hit by the 2013 flood, had been affected in 2002 or later, this holds particularly for Thuringia, Lower Saxony and Baden-Wurttemberg. The number of surveyed residents from these regions was, however, lower than aimed at. The correction of flood experience will decrease the flash flood data to a great extent, increasing analysis uncertainty. However, we will analyse the indicator "threat appraisal" with regard to corrected data (all households that experienced a flood for the first time), discuss the results and put a figure (figure D) in the appendix.

**Reviewer quote, paragraph 2, 6, 7:**

Specific comments: Page 4, line 8-20, these are some very technical sentences, could you explain your approach in a more intuitive way and introduce the technical methods later. Currently this is difficult to read without prior knowledge about the statistical methods that are applied.

2.5 Explain why you use Bayesian statistics, you now just jump into the explanation without first motivating the choice.

2.5 Why did you choose to use Bayesian statistics if no prior is available? What is the advantage of using Bayesian statistics?

**Answer to paragraph 2, 6, 7:**

Thank you for this suggestion. We will rewrite sentences that are too technical and give explanations of the statistical methods beforehand. We will further elaborate the choice for Bayesian statistics. We decided to include Bayesian statistics without specific prior information since any valuable results from other studies could theoretically be integrated into our analysis in future. Suggested changes in the text:

> "Secondly, the differences in the indicator distributions, i.e. shifts to lower or higher indicator ratings, are assessed for each flood type. To answer the second and third hypotheses, a "planned precaution" indicator is created first. In a next step, the Bayesian approach and negative binomial regressions are applied and resulting probability distributions of conditional variable dependences as well as regression coefficients are evaluated. The Bayesian approach has been frequently used in psychology (e.g. Wetzels et al., 2011) and other disciplines. It assesses the data uncertainty which is particularly helpful among studies that rely on relatively small data sets, while prior information independent of the data can be included (Van de Schoot et al., 2015). Since this study relies on small data sets, using the Bayesian approach as a supportive analysis helps to interpret main results. By revealing data and model uncertainties, the reliability of future prediction models that are based on these data sets can be evaluated in advance. Accordingly, this study considers Bayesian inference as a method to assess variable relations, that are based on conditional probabilities and related uncertainties. Preliminary assumptions such as e.g. linear variable coherences are therefore not required. ~~Furthermore, this approach evaluates the specific variable applicability for a potential prediction of a response variable, in this case the "planned precaution" indicator.~~ Bayesian statistics were also chosen due to the fact that the method enables prior knowledge to be taken into account, for example in following studies that use similar Bayesian approaches."

**Reviewer quote, paragraph 3:**

Page 4 line 24-30: Could you sketch in a bit more detail how you see this being used in the future. We don't know these psychological indicators for everyone when we make a damage model. It might even be easier to ask directly about

precautionary measures than to assess their psychology. Using social media information as proxy might be a solution but I like to see these arguments made a bit more thoughtful and if that's the way to apply it I like to see that back in the discussion and maybe a recommendation to study how social media clues can be linked to the indicators used in this paper. You mention several times the "protection motivation theory", please give a brief explanation of this, you can't assume all your readers know about this.

**Answer to paragraph 3:**

Thank you for the comment, we will elaborate on the topic of alternative data sources as well as new approaches to gather valuable data in the discussion. We will further give a brief outline of the Protection Motivation Theory beforehand. Suggested changes in the text:

"In this context, the protection motivation theory (PMT) (Rogers, 1975) has been frequently used as a psychological model to explain the risk-reducing/protective behaviour of affected individuals by analysing the influencing factors on coping strategies and potential positive or negative responses. Originally evolved in the health sector, the PMT gained attention in the domain of natural hazards over the years (Mulilis and Lippa, 1990; Grothmann and Reusswig, 2006; Bubeck et al., 2017). The model relies on two main cognitive processes - "threat appraisal" and "coping appraisal" – to describe the mental response to a specific threat. Threat appraisal is composed of the perceived consequences and probability of an event. Coping appraisal comprises the variables "self-efficacy" (perception of how well a person is able to carry out protection measures), "response efficacy" (how effective the measures are believed to be) and "response cost" (the perceived cost in terms of money and effort) (Rogers, 1975; Bubeck et al., 2012)."

(…)

"• A better understanding of this connection might help to improve future vulnerability and risk estimations and may facilitate the use of alternative data sources to estimate the state of individual precaution. For example, data from online surveys, social media and communication platforms offers a lot of potential to assess individual mental coping strategies such as evasive behaviour or active remembering after severe events. With the help of advanced intelligent learning algorithms (e.g. random forests, neural networks and deep learning), psychological profiles could thus be created. Those might be used to develop sophisticated models and predict the state of precaution in areas which have not been flooded recently, all based on data given voluntarily by residents. Surveys that capture the state of precaution are still an alternative option."

(…)

"An issue of telephone surveys is that the data is becoming biased towards older participants when based on landlines (Greenberg and Weiner 2014). Alternatively, by implementing and making use of online surveys, smartphone applications and contracts with companies, valuable data could be collected accounting for people from all age groups. For further use, algorithms such as Neural Networks or deep learning algorithms may be applied on this data to create

or categorize psychological aspects such as the expected level of burden or evasion in case of an event. Those techniques might result in good predictions of psychological behaviour and the connected precaution motivation and can theoretically be transferred to other regions but yet imply certain challenges. Firstly, large amounts of consistent and high quality data have to be collected on condition that data security and personal rights are considered. Secondly, the interpretation of results in terms of causality and meaning is hampered due to the black box character of the analysis, even though potential results might show a certain robustness."

**Reviewer quote, paragraph 4 & 5:**

Page 7, line 5, please first explain what burden and evasion are before explaining the motivation behind it.

Give a proper explanation of Kruskal-Wallis rank sum test, Dunn‘s Test, the Jensen-Shannon divergence and regression tests directly after you first mention these methods. Maybe don’t mention them too early in the text. Give both an intuitive and a brief mathematical explanation of the methods.

**Answer to paragraph 4 & 5:**

As you suggested, we will describe and explain the methods as well as key terms beforehand and in a clearer way. Suggested changes/additions in the text:

"Subsequently, the four main indicators are defined as “threat appraisal”, “coping appraisal”, “burden” and “evasion”, which also show low intercorrelations and offer a certain comparability to other studies. The four indicators are thus defined and created as follows.

According to the PMT, threat appraisal consists of the perceived probability of being affected again by a flood event and the perceived impact of such a future event. Coping appraisal comprises self-efficacy, response efficacy and response cost which describes the self-rated ability to implement a protective measure, the perceived efficiency of a protective measure and the perceived cost of the protective measure, respectively (Grothmann and Reusswig, 2006; Bubeck et al., 2012).

Burden describes a measure for the negative psychological load of the experience and consists of the single variables "often thinking of the event" and "stress still today". Evasion comprises the variables "avoidance" and "fatalism" and can be seen as a measure for the effort to get the experience out of one's mind for various reasons. Burden and evasion were developed by following the general procedure in psychology surveys to combine expressive psychological items (e.g. Ware and Sherbourne, 1992; Kroenke et al., 2001) and taking high correlations among psychological variables into account."

(…)

"The distributions of threat appraisal, coping appraisal, burden and evasion were further analysed using the Dunn's Test, which is based on the non-parametric Kruskal-Wallis rank sum test results. These tests are suitable for assessing the differences among the distributions of ordinal-scaled data, which does not fulfil assumptions of normality and equality of variance. Here, the Kruskal-Wallis rank sum test is preliminary to the Dunn's Test and calculates discrepancies among the rank sums of all values within the compared indicators. The derived Kruskal-Wallis statistic is then compared to the expected average difference among the sum of ranks via Dunn's Test. Similar to a power analysis, the effect size and significance are revealed for a given sample size. The outcome represents a measure for the disparity and shift of compared distributions. This approach reveals significant differences in psychological impacts which were predominantly caused by weak flash floods, strong flash floods and river floods."

Regarding the Jenson-Shannon divergence and the negative binomial regression, the respective parts in the text will be converted in a way that the explanation follows directly after mentioning the methods for the first time.

**Reviewer quote, paragraph 8:**

Figure 2: why is threat appraisal lowest for strong flashfloods? Does it make sense that if something extreme happened to you, you feel the probability that it will happen again to be lower? (your argument on page 11, line 10). Maybe threat appraisal is lower because they only experienced it once while the river floods and weak flash floods were experienced more frequently. If however you would go to another region where only one weak flash flood or river flood was experienced these results may look very different. You should probably discuss that limitation in the study.

**Answer to paragraph 8:**

We agree with the statement of the reviewer that this means a certain limitation of the study. Therefore, we will add another analysis in the appendix with the corrected data in terms of flood experience for the indicator "threat appraisal" (see also Answer to paragraph 1). The limitation will be discussed in a more elaborate way, yet we believe that our general statement in that case ('it has been such an extreme event that people perceive it as unlikely to happen again') holds true. Suggested changes in the text:

"Although Hopkins and Warburton (2015) showed that flash flood experience does not necessarily lead to higher risk perceptions, it is unknown, to which degree lower feelings of threat are caused by a lower flash flood experience itself. Since almost all surveyed households experienced a strong flash flood for the first time (82%), they may not believe to be affected again. However, an analysis of threat appraisal with corrected data in terms of flood experience (all households that experienced a flood for the first time) reveals a similar picture, i.e. threat appraisal is significantly

lower for people who were affected by a strong flash flood in comparison to people who were affected by weak flash floods and river floods (see appendix, Figure D). This again supports the findings of Hopkins and Warburton (2015)."

**Reviewer quote, paragraph 9 & 10:**

Figure 4: This figure is not very intuitive can you explain a bit more what the reader sees here.

Figure 4: Why do you see the double peaks in the probability distributions?

**Answer to paragraph 9 & 10:**

Thank you for the hint. Firstly, we refer to the explanation graphic, Figure 1, where we will add more details to the description. We will also describe the method in a better way, changing the respective text passage underneath Figure 4. Suggested changes in the text:

"Figure 1: Example graphic explaining the creation of the weighted arithmetic mean posterior. The double peaks are a result of the combination of all posteriors in one plot that are calculated for each variable combination. The posteriors are weighted according to the sum of occurrences within the dataset. In this case the weighted mean posterior means that, given the example dataset of 20 data points, it is most likely that a specific predictor variable rating occurs together with only one specific response variable rating to 80%."

(…)

"The weighted arithmetic means of all posterior distributions reveal in general a wide range of likely probabilities for the conditional dependence of variable ratings. In the case of weak flash floods for example, it is second most likely (second highest posterior peak) that a particular burden rating is always reported together with a specific rating of the planned precaution to 52 per cent (most likely to 9 per cent due to the highest posterior peak at this point). For coping appraisal, the most likely percentage would be 7 per cent. For threat appraisal and evasion, the most likely percentages are 10 and 19 per cent, respectively (Figure 4, top left)."

**Reviewer quote, paragraph 9 & 10:**

Page 14, line 17: You say this is common practice in psychology, can you provide a reference for this?

**Answer to paragraph 9 & 10:**

A reference will be added. Suggested changes in the text:

"When comparing the analysis of the psychological indicators and the single variables, it can be summarised that a combination of items, as it is practised by e.g. Ware and Sherbourne (1992) and Bei et al. (2013), does not lead to more consistent and meaningful results in this case which is mainly reflected by similar JSDs."

**Reviewer quote, paragraph 9 & 10:**

Page 15, H2: I think the findings make much more sense than the hypothesis.

**Answer to paragraph 9 & 10:**

We think that the Hypothesis is justified since personal experience and conversations with flood affected residents indicated a high level of burden after a severe flash flood event, which could also lead to negative responses and low motivation deal with any aspects and implications of the flood event again. Still, we believe that these are interesting negative results which support other studies such as Bei et al. (2013), who reported that affected people with worse mental and physical health show a higher willingness for coping strategies.