Response to Reviewer #1

The authors made an extensive revision of the manuscript and have improved many aspects. Very helpful are the tables describing terminology and variables. Unfortunately, the most important weakness of the study (overfitting) was not addressed in the revised version. In my original review I asked if overfitting can be ruled out. The authors admitted that this is not the case but no consequences followed.
Overfitting is addressed in the newly added Appendix A of the paper where we present results of simpler versions of the model and show that the magnitude of the results do not change if a simpler model is selected. However, the simpler models are severely prone to omitted variable bias. Therefore we have decided to keep the original model in the main text because they take other important confounding factors into account and are, in our view, the most appropriate for understanding the relationship between heavy rain and natural hazards along the rails.

An easy to read publication about overfitting can be found under the following DOI:
10.1097/01.psy.0000127692.23278.a9
Citing from this publications:
1) "Taken to its extreme, if the number of unknowns in a model is equal to the number of observations, the model will always fit the sample data perfectly, even if all the predictors are noise, ie, entirely unrelated to the response variable."
2) "For linear models, such as multiple regression, a minimum of 10 to 15 observations per predictor variable will generally allow good estimates."
3) "In the case of models with a binary response, if the number of events is smaller than the number of nonevents, the limiting sample size is the number of events."
4) "If we cannot gather a sample of sufficient size, we have to find ways to simplify our model..."
We have opted for reducing the complexity of the model (option 4) and we show in the appendix that the primary interpretations hold with a less-complex model.

In the light of these statements it is absolutely essential that the authors perform tests to demonstrate that the results (e.g. the statistical significance values) are meaningful and not artefacts of overfitting. Such a test could be for example a cross validation. The prediction for the years not used for training should have more skill than forecasting the climatological probability of the respective season. The large number of extra coefficients associated with the random variable, alone, makes me wonder if there can be any skill in the statistical models - especially the ones for gravitational mass movements.
We would like to emphasize that the main goal of the paper is to shed light on the relationship between heavy rain and the three natural hazard processes, and not to develop an accurate predictive model of natural hazards. Therefore we have determined that a cross validation procedure, which evaluates the models based on their predictive capacity, is not the appropriate approach to tackle the issue of overfitting for this analysis. Instead we have chosen, as mentioned, to present simpler models with different permutations of the control variables, and demonstrate that in cases that are at most risk of overfitting, selecting a simpler model will not alter the conclusions.

In the revised manuscript the authors write for example: "Despite the lower number of observations of events available for the gravitational mass movements, the model and chosen variables describe the relationship between gravitational mass movements and heavy rainfall events accurately. In the case of tree fall events the actually by a magnitude higher number of data points available for the

calculation does not lead to a better model fit...."
I believe that the perfect fit for gravitational mass movements is an artefact of overfitting.
For gravitational mass movements, the odds ratios of heavy rain at the mean values of the meteorological variables (which can be found at the bottom of Table A3, columns (4) – (6)) are similar in magnitude to the odds ratio of heavy rain in the simple logit regression in Table A3 column (1). If the simpler model were to be selected, the magnitude of the effect of heavy rain on gravitational mass movements would change little. Nevertheless, the loss of statistical significance in the coefficient of heavy rain in columns (4) to (6) of Table A3 makes it difficult to make a solid conclusion based on the current data alone. More data is required for more robust results.

Something that should also be improved before publication is the description of the statistical models.
I still find it very confusing to understand what the final statistical model looks like. In the answer to the review the authors state: "For each hazard model, 5 coefficients and one random variable had to be fitted." This agrees with equation 2 as $\boldsymbol{\beta}'\mathbf{2}$ includes 3 coefficients.
Later in the manuscript the authors write: "Annual and seasonal dummies are also included to account for the fact that the number of natural hazards varies greatly in different years and seasons." These additional variables don't show up in the equations and are not included in table 2. The hazard indication for slope and embankment landslides is also only mentioned in the text but missing in table 2 and the model equations and/or the vectors listing the control and random variables. With these extra variables the model becomes more complex and includes even more coefficients.
To allow the reader to understand the study, the full model equations and a complete table with all the variables is essential. The incomplete tables and equations make this part untransparent.
We have adjusted the description of the statistical models in Section 2.2.3. to explicitly describe all included control variables and interaction terms. We have also adjusted the calculation of the odds ratio to reflect the fact that the odds ratio will vary depending on the values of the meteorological control variables due to the interaction terms. In light of this change in the calculation, we have added Table 3 to the results section presenting the calculated odds ratios at the mean and median values of the meteorological variables. In the Appendix, the full regression tables area also presented will all included variables, seasonal and year controls and interaction terms.

Outline: In their reply to the reviews the authors explain why they mixed the contents of sections. I must admit I am not too happy with this decision. When reviewing the manuscript I found it difficult to find the relevant information when trying to retrace certain aspects. The data section includes results (the analysis of the seasonal variation of the combined hazard). The rail segments are provided by DB (and not work done for this study). Therefore they should be described in the data section and not in the methods section (L272-281). It is up to the editor to decide if the outline can stay as it is or if it needs to be revised.^
There was no mention by the editor as to whether a change in the outline is necessary.

Minor remark:
Equation 3: x needs to be replaces with epsilon
The greek letter epsilon is not used in any of our equations.