

Interactive comment on "Modelling Vulnerability to Severe Weather" by Tobias Pardowitz

Anonymous Referee #1

Received and published: 11 October 2017

The concept of the work is interesting; however there are several problems related to the methodology, the presentation of the results, the contribution of the resulting model, and consistency within the text. Furthermore, special attention to English language has to be paid, as there are several sentences that need to be corrected or rephrased.

Major issues: 1) A major issue is the statistical methods applied and the presentation of statistical results. Nowhere in the text or Tables are any notes provided regarding the calculated correlation coefficient. What kind of correlation (pearson's r, spearman's rho...??), are data normally distributed, any tests made for that, what are the p-values for testing significance of results. P-values are actually very important, along with the relevant significance limits (e.g. <0.05, most commonly used). The same for regression results, for which tables must provide also the standard errors, whether the coefficients are standardized, and what is the overall significance of the model (this I assume is high judging from the R2). Also, the number of observations for each correlation and

C1

regression must be provided.

2) Important methodological deficiency is related to:

a) Failing to address the exact weather cause of operations and losses. The operations alert keywords do not necessarily show the weather cause. This could be done only by using meteorological stations near the operations and include meteorological parameters in the analysis because these are the hazards.

b) Lack of flood-specific operation or insurance data. Floods constitute the most damaging and expensive weather-related data. I would expect this to be clarified. Floods can cause damage to buildings, cars, trees, displacements, thus the model would have a serious value. In addition, it is not clear whether loss data have been identified as to the weather-related cause. It seems that loss data are not distributed according to the cause(s), while only windstorms/thunderstorms are addressed.

3. The actual contribution of the models is pending, since the predictor factors chosen ('...topographic features, land use, urban structure) are more or less constant throughout the examined period, and depict only the structure of the city. Predictors would have been the meteorological hazard parameters, which are missing from the models. I suggest the authors identify the daily weather conditions, based on meteorological data from weather stations in the city. Otherwise, the models only depict what has happened in the past based on an unchangeable 'environment'. Prediction based only on stable (endogenous) factors is not a prediction. Even the season parameter, for which a discussion was made in the methodology, is not included in the models as a dummy variable. Out of the two objectives set by the authors in the introduction (p2, lines 24-26), only the second is to my opinion reached: the identification of hotspots for weather impacts. It is not possible to 'predict the local occurrence rate of operations (p. 10, I13)' based only on these vulnerability indicators and on regressions of annual values.

4. Following the previous comment, there is a high correlation between the predictor

factors and the outcomes. Specifically, I find it logical for the road network density to be highly correlated with traffic problems, or housing density with construction-related impacts. This may be a methodological problem for multiple regressions. However, I can't evaluate the regressions without the required additional statistical information.

5. The authors use the expression 'potential predictors'. Then, it must be noted exactly what is the methodological contribution, because a serious problem for scholars is to find suitable and reliable predictors for relevant studies. Again, the problem is that meteorological variables or other hazard-related variables are not included or discussed in the present analysis.

3) Other major issues

Section 2.1. Are flood-related damages included? Please specify, since floods are the costliest weather-related hazards.

Section 2.2. Berlin-wide damages are available on daily basis for the period 1997-2011, while data on zip code level (190 within Berlin) is available for selected events only: how many events? In general, this paragraph is not clear. If only wind and light-ning are included, how can the comparison with operations be made? The selection of 5 events was based on which criteria?

P.4. line 1: evaluated for what exactly? Relating to the spatial distribution of operations? And why in section 4.1 you mention 6 events instead of 5?

P.4. line 5: the second part of the sentence – regarding hail- is based on other references? If it is so, consider separating the sentence, because the first part refers to the analysis of losses, thus it is part of the work done by the authors.

Section 3.1. The comparison of operations and loss data is strange. As explained in section 2.2, loss data refer only to wind and thunderstorms, while operations may be related to many other causes. So, what is correlated? It is confusing.

Section 4.1. In general, there is an inconsistency in these correlations since flood-

C3

related losses are not included. Many assumptions have been made in the interpretation of results for this reason. E.g.: P7,I2 'coinciding with roofing damages or other wind-caused building damages' Is this an assumption? Since losses do not include flood-related damages, correlations for the water-related categories are not useful. Operations and losses are very likely to refer to different hazards.

Other methodological issues:

In the Abstract the building density is suggested to be the most influential predictor. However, results suggest that building coverage has the strongest effect.

The expression 'water-related' operation refers to what specifically? Other than flooding, or flooding is included?

I would suggest you explain in the text what 'operations are equally distributed over winter and summer half year' mean. E.g. in a parenthesis (e.g. May to October, for summer half). Also, some local weather info would be helpful. For example, during summer half-year is there more wind and storms? And why do you make this 2-seasons choice?

Why is population density not included in the predictors?

What about the validation results? Why they are not shown?

References in the introduction regarding the use of fire service operations as impact indicator are missing. However, there are relevant papers.

What does a correlation between predictors and the TOTAL number of operations show, since damages can be due to different/various causes?

Minor: Please, follow the same citation format (e.g. see differences in page 2, lines 5 and then 7.

Please, follow the same term for fire brigade operations. Are they callouts or operations? The authors use different terminology. However, callouts and final operations usually are 2 different things. In page 8, scetion4.2, the term used is 'operation densities'. However, it is still the number used.

P.4. line 2: is it storm or thunderstorm specifically?

Sections 3.2, 4.2. Consider improving the title of this section. Correlation of What (?) with potential vulnerability predictors.

P9, I12: 'Additional predictor variables from the CORINE land cover dataset are assessed'. Only one is shown in the discussed picture (artificial surfaces)

P10.I9: Consider removing 'considerable'. This is to be evaluated in the regression results.

Sentences that should be definitely corrected/rephrased (not all cases are included. The article needs to be carefully reviewed for grammar):

P2, lines 12-13.

P3, lines 14-15.

P4. Lines 1-2

P4. Lines 14-15 (consider changing the beginning of the sentence)

P7. Lines 22-23. A verb is missing.

P7, I6: 'Correlating tree-related and water-related operations'. Consider correcting as: Correlating tree-related WITH water-related operations?

P8, lines 3-4. A verb is missing.

P8, I5: Consider using 'events' instead of 'areas'

P8, lines 11-12, 17-20.

p10.l2: Consider deleting 'are particularly vulnerable. The sentence needs rephrasing.

C5

p10, l29-31.

P11, l11-13.

Errors (not all errors are identified here. The article should be carefully reviewed for language):

P2. Line 18: Consider replacing 'is' with 'are'. Consider deleting 'compare' in the parentheses that refer to the Table.

P3: Line 16: 'which naturally occur exclusively in winter: the phrase is repeated after 4 sentences. Consider deleting one of these.

P.4 Line 24: ha corresponds to 'sq. m', not 'm'

Section 2.2. P.4. line 1: Consider correcting: correlation 'with' something.

p.5 line 7: consider correcting as: 'are analysed'.

p.6, line3: 'are; is missing before 'addressed'.

p.6 line 17: consider correcting ' the investigation area berlin'

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-326, 2017.