

Interactive comment on “Avalanche risk in backcountry terrain based on usage frequency and accident data” by F. Techel et al.

Anonymous Referee #1

Received and published: 10 September 2014

General comments: The paper with the title: “Avalanche risk in backcountry terrain based on usage frequency and accident data” is of general interest to the NHESS reader community. The authors base their analysis on data available on two web portals in combination with snowpack and avalanche accident data. The aim is to include usage frequency in the analysis in order to improve risk assessments. This is certainly an approach worthwhile. However, the title promises a more detailed risk analysis than is presented in the paper. The regional analysis includes new aspects, however it is questionable whether the data and used statistics allow such a regionalisation. Anyhow several interesting aspects have been discussed. I understand that it is extremely difficult to deal with data that are prone to great uncertainties and find the appropriate statistics. In most parts it is not clear what methodology and outcomes are intended to

C2038

be achieved or how they should be interpreted. In general the used data is suitable for an analysis, however the data has to be discussed in regard to the use of appropriate statistics. As there is no clear structure in the paper the authors switch between statistics and it is difficult for the reader to follow the working steps. There are studies on how to deal with different data types and data quality that could be used to improve the analysis of this paper. Previous work has been credited, mostly to other Swiss researchers, however even though the work from other countries is partly mentioned a detailed interpretation is missing. Some of the technical terms seem to be awkward and the paper should be read by an editor for grammar and wording. In the following I made some comments by heading in order to underline the above statements.

Introduction The introduction picks up various aspects on avalanche accidents and the analysis of data. However, what is missing is a clear introduction on the motivation and the exact goals of the paper. The aims are formulated in a very general way and are not related to the literature review above. Even though risk is explicitly mentioned in the title a definition on how risk is defined in the paper is missing. E.g. the term collective risk differs in meaning from other risk studies. The introduction is not very well structured and it would be desirable to restructure with emphasis on relevance to the topic or specific aims.

Data The data is more or less listed without informing the reader why it is needed to meet the aims of the study. An introductory sentence would be helpful. The description of the data is not consistent and it is often difficult to understand how and why it is used in the subsequent analysis. Sometimes the total number of observations is given, sometimes not, sometimes percentages are given, sometimes ratios etc. Please reconsider the presentation of the data (all variables should be explained in a similar way) and point out what is really important for the analysis. The data description should include e.g. the total number of observations by variable, the numbers by class, how you standardized them and why etc.. Up to now it is not possible to follow the analysis steps, e.g. when and how you standardize your data. In the following I would like to

C2039

point out some aspects that were unclear to me: - Activity: who is allowed to register? How do you interpret the entries regarding uncertainties? In order to understand why you rely on the data a better description of the possible inputs in the database would be interesting. In Table 1 you say what has been reported to 100%, but we do not know how users report activities. E.g. is the summit elevation entered by the users or is there already a form you can select the summits from? How high are the uncertainties in each entry? What is subjective, what objective? How is the route difficulty determined?... - Accidents: I cannot follow your argumentation why you use 10 years of avalanche data and how you link it to 5 years of frequency data? What is your reasoning to exclude off-piste accidents? ... - Weather: the classes are ok, however how many days are in each class you consider? What are the numbers by region and category? How many observations are available for each region and how well do these represent the entire area? ... - Avalanche danger: 17:00 from the day before? Valid for the day of interpretation. - Snowpack: Why do you exclude south slopes? Or is there no data available? How many observations are available and how are they regionally distributed? Do you only consider snow observations from the same day or do you extrapolate them over time? If so why and how? How many years of data do you consider?...

Statistics The description of methods and statistics is poor. In a first step the kind of available data should be analysed and explained why the subsequent statistics can be applied to the available data. The way of reporting frequencies, ratios, significance and correlations are not consistent. It is e.g. difficult to reconstruct some values. E.g. in Table 4 the day of the week is reported as percentages, however the reference in the paper is made using ratios. Now the reader has to divide the percentages by 5 (week days) and 2 (week end) and then the ratio and Table fit. This is only one example. You use modal values for activity and accidents – how? While a modal value for the weather data is understandable, I do not understand what kind of modal value you use for the climate regions where users recreate mostly (p.5123). You use X^2 , Spearman rank correlation, Mann-Whitney, populations, frequency data, contingency tables...

C2040

However, in your results and conclusion you only report p-values, but it is not possible to know what kind of test you are referring to. What is your population? ... Please be more concise in your description and reporting of the statistics. Why did you come up with an interpretation of P-values ≤ 0.1 to be marginally significant? Is that to account for the great uncertainties in your data? Why did you reduce the level of significance otherwise?

Results and discussion Again I believe that a better structure would help the reader to understand your interpretation of the results, especially if you focus on your key aims and do not try to subjectively assess patterns in demographic data. Your description of the activity data can include aspects on how many male/female users enter data etc., however a direct link to the avalanche accidents seems not appropriate. What is your statistical hypothesis behind this statement/interpretation? The discussion should be rewritten once the statistics are better explained and be reviewed again.

Comments on Tables and Figures. Tables and Figures should help the reader to get a good overview of the most important aspects of the paper. However, I found some of the tables and figures are missing real important information and that the headings are sometimes misleading. In the following I pointed out some specific remarks: Table 1: Do you consider the Table as meaningful? Is there not more interesting data that should be summarized in this table? Table 2: why do you only report the total number and not also the numbers by group? Table 3: I do not understand the importance of the table and of the interpretation. Figure 2. a) What am I supposed to see? The modal value is the value that occurs most often in your data. Here it is a percentage? What are you comparing it to? What are the total numbers behind the percentages? 2. c), d) the descriptions in the text and what I see in the figures do not match.

Are all of the figures necessary?

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 2, 5113, 2014.

C2041