Santiago, Chile, June 2016

Dear Editor,

My two co-authors and I acknowledge the two reviewers for their work and their valuable comments. They have allowed us to significantly improve the quality of our paper. In the following we present a response to their comments. If we are allowed to submit a revised manuscript, we also will include in the final document the specific comments made in the *pdf* document attached by the reviewer #1.

On a general way, it seems that for both reviewers, the objective of the paper is unclear. The objective of the study is to consider different methods of rockfall hazard assessment, and evaluate the differences in terms of levels of rockfall hazard. The levels of rockfall hazard considered in the study correspond to classical rockfall hazard levels: *very low*, *low*, *moderate*, *high* (Copons et al, 2008; Bauer, 2011; OFEFP, 1997). These levels are commonly used to build hazard maps, for risk management in urban areas. Thus, it is possible to compare these levels, obtained using different methods. Moreover, the influence of the level of expertise on the result, and so on the obtained hazard levels, is also investigated. The objective of the paper is not to evaluate the "true" level of rockfall hazard, but to compare the evaluation process, considering different levels of expertise and different methods. This general objective has been clarified in the introduction to the paper. Moreover, the comments of both reviewers helped us to improve the clarity of the entire paper.

Reviewer #1 and #2:

Both reviewers made comments regarding the concepts and vocabulary used in the paper. In particular, according to them, the concepts of "probability" and "susceptibility" should be clarified.

Reviewer #1: A major point of review as far as it concerns the methodological concepts used in this work has to do with the definition of the probability of a rockfall. It is not clearly described in the paper whether this term refers to a spatial or temporal probability of occurrence. The interpretation of the probability as a term depends on the uncertainties that are taken into account for its definition. As a result, probabilities referring to different types of uncertainties cannot be compared, because they represent different values. This point is not clear in this paper and probability concepts are mixed.

Reviewer 2: Several descriptors are used for the temporal likelihood of rockfall failures (occurrence probability, temporal probability, susceptibility to failure and failure probability, Tables 1 to 7). This is quite confusing: ***a)*** *Susceptibility*

is assessed in the manuscript using both spatial conditioning factors and those related to (potential) temporal occurrence. This does not follow internationally accepted standards (e.g. as the defined by JTC-1 in Fell et al., 2008), in which susceptibility corresponds to the volume and spatial distribution (i.e. size and spatial probability) of potential landslides. Although it is expected that landsliding will occur more frequently in the most susceptible areas, in the susceptibility analysis, time frame is explicitly not taken into account (Fell et al., 2008). ***b)*** *In page 9, occurrence probability and temporal probability for applying LPC method are defined as addressing the questions if a rockfall will occur (will the rockfall occur?) and when will the rockfall occur (as the annual frequency of occurrence), respectively. The difference is not clear to me. They provide the same type of information, because a null temporal frequency means a null occurrence probability (the rock fall will not occur). The words used for describing the "temporal probability" (imminent, very short term) does not suggest annual probability but time expected to failure, which is a different random variable. Could you clarify this issue? Time to failure is more difficult to be estimated than annual probability even when monitoring data are available.* ***c)*** *Use of failure probability for a hazard matrix (Tables 5 and 7) can be misleading. A high failure probability means a high temporal probability of occurrence, which typically lowers when landslide size increases. Hazard level increases with both temporal probability and landslide size.*

Response:

Indeed, the concepts, as presented in the paper, may seem unclear. In the current version of the paper, we use the following concepts:

- "rockfall failure probability" (line 21, page 2), which refers to the probability that a rock became unstable;
- "occurrence probability" (LPC method), which answers the question "will the rockfall occur?";
- "temporal probability" (LPC method), which answers the question "when will the rockfall occur?";
- "susceptibility to instability" (LPC method), which is the combination of the occurrence probability and the temporal probability;
- "occurrence probability" (SMR-based method), which is a combination of the SMR value and the site activity.

To simplify and clarify the reading of the paper, we propose to modify the vocabulary. As presented by Fell *et al.* (2008), the hazard is a condition with the potential for causing an undesirable consequence. The description of landslide hazard should include the location, volume, classification and velocity of the potential landslides and any resultant detached material, and the probability of their occurrence within a given period of time.

Based on this definition (Fell *et al.* (2008)):

- the term "***rockfall hazard***" is used when referring to the level of hazard (previously called "rockfall failure probability"). It corresponds to the

combination of the occurrence probability, the temporal probability and the intensity (volume). We propose to define this vocabulary at the end of the introduction;

- the terms "***occurrence probability***" and "***temporal probability***" are not modified into the paper. However, their definitions are modified. The occurrence probability refers now to the question "can the rockfall occur?", and not "will the rockfall occur?". Moreover, as noted by reviewer #2, the temporal probability is defined as the time expected to failure. In other words, the temporal probability is defined by the probability that the failure occurs before an expected delay. It is a parameter particularly hard to evaluate, even when monitoring data are available. However, it is a key parameter in the evaluation of rockfall hazard. We decide to introduce this term in the same way as in the LPC method guidelines (Laboratoire des Ponts et Chaussées, 2004 – in French), as presented in Table 2.;
- the term "susceptibility" is removed, as it is not coherent with the definition proposed by Fell et al (2008). To replace it, the term "***predisposition to instability***" is used. It corresponds to the combination of: (1) the *occurrence probability* and the *temporal probability*, when dealing with the LPC method, and (2) the SMR value and the site activity, when dealing with the SMR-based method. Then, the *predisposition to the instability* is coupled to the volume to evaluate the *hazard level* (the news terms will be used in Table 5 and 7). Using this terminology, it is clearer that both rockfall hazards assessed using the LPC method and the SRM-based method can be compared, because is compared a combination of the *predisposition to the instability* and the volume in both cases.

Please note that in the study proposed here, the spatial probability is not evaluated. This choice has been made as initial assumption to simplify the analysis, in particular for the students. We believe that it does not remove the relevance of the study.

To go further, we propose to modify the current title into "Influence of expertise on rockfall hazard assessment using empirical methods". This title clarifies the fact that the objective of the paper is not to determine the "true" rockfall hazard, but to investigate the evaluation process considering two different methods, and different levels of expertise. The comparison realized in this study does not allow the best method to be highlighted. It leads to a better understanding of the rockfall hazard assessment process with different levels of expertise, considering two different approaches. This point will be discussed further in the discussion part of the paper.

Reviewer #1:

"The quantification of the probabilities misses explanation."

Response:

In the paper, the probabilities are not quantified for any of the methods. In the LPC method, the parameters and the probabilities are not quantified. In the SMR-based method, only the SMR value is evaluated and quantified. However, to evaluate the predisposition to instability, the SMR value is combined to the activity level, which is not quantified.

In part 3 "Results", the level of rockfall hazard is coded (from 1 to 3). This quantification of the values is used for the statistical analysis. This point will be better explained in part 3 "Results", lines 13 to 18.

"*An extensive state of the art on the topic of the calculation of rockfall probability is missing. I believe that before proceeding with some specific points, a general review of the methodological concepts should be made respectively. The methodologies and the results should be better explained and some of the conclusions at the discussion should be checked to make sure that they are coherent with the results.* "

Response:

As explained previously, the concepts used in the paper have been clarified. A new definition of the hazard is provided, and the entire paper is now based on this new definition. The different concepts are presented Figure 1. We believe that a better presentation of the concepts will help the reader to better understand the results themselves. Moreover, a better explanation of the objective of the paper will also help for that.

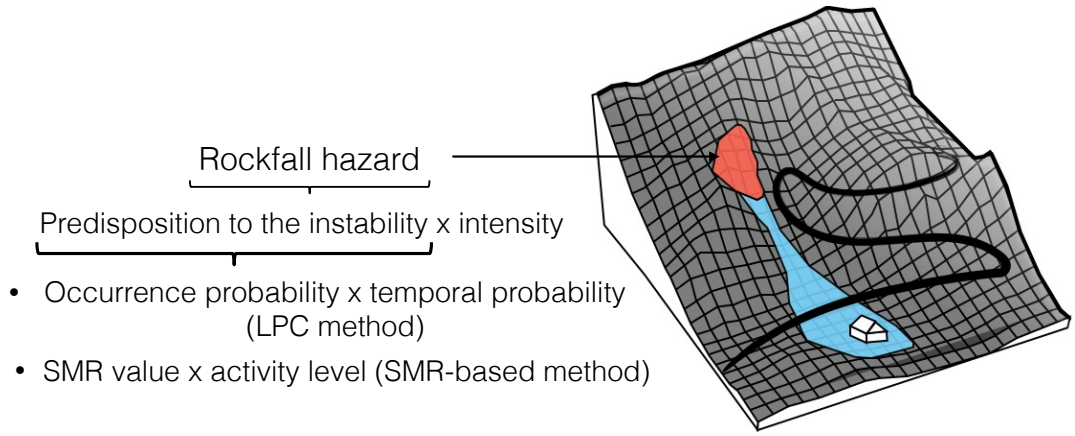

Figure 1. Concepts used in the paper

"The English language would need a thorough review as well. "

Response:

The paper has been reviewed by the American Journal Expert (AJE) support team to check the English quality of the paper before submission. However, we have carefully reviewed the language of the paper one more time.

Reviewer #2:

1) A good site characterization seems essential in order to allow readers to carry out your own appraisal of the hazard and evaluate the interest of the study. Nevertheless, a very basic description of the site and sectors is provided. Especially important are data on joint orientation, prevailing failure mechanism and past rockfall occurrence. Is traction failure of overhanging blocks the main mechanism in the site? This is a key issue, because SMR method is routinely used to assess rockfall susceptibility of failure mechanisms involving shear on joints (including flexural toppling) tough it seems not applicable to failures caused by traction.

Response:

Indeed, in the past version of the paper, there was not enough information related to the study site. The authors agree that more information is valuable to the reader. The following information have been added to the paper:

- The prevailing failure mechanism on the three sectors (after Hantz *et al*., 2003). For sector 1, the main failure mechanism is column toppling. For sector 2, it is wedge slide. For sector 3, it is overhang failure; the latter type of failure is also present on the two others sectors. The overhang failure is the result of a traction failure for some of the small blocks at the top of the three sectors. However, there also exists a failure mechanism that involves shearing of joints on the studied 3 sectors. A comment related to this has been added to the discussion part (page 17, lines 19 to 27).
- Past rockfall occurrences: the test site has been chosen because of its history. It used to be a climbing site, and was closed after repetitive rockfalls. However, we don't have any quantitative information concerning the occurrence of the phenomenon. It is not a monitored site. A comment has been added to part 4.2 "Discussion" concerning the choice of the test site (page 17, lines 19 to 27). Moreover, when presenting the test site, we explain that "the site has been previously classified as presenting a high-level rockfall hazard (Moiriat et al., 2008)" (lines 22, page 3). We explain that the methodology used to evaluate the hazard is the following: (1) bibliographic review of all the documents available, and (2) field recognitions. The first step leads to evaluate the risk area at a departmental scale. The second step leads to prioritize the level of rockfall hazard on the area of study.

No systematic information concerning the joint orientations is available.

3) Information or qualitative criteria used for estimating occurrence/temporal probability of Tables 1 and 2 should be specified in order to

make the results obtained by different people comparable. Similarly, geomorphic indicators used in the SMR-based method should be also described in the manuscript. Quoting El-Shayeb (1999) work seems not sufficient. What type of morphological traces where found in the test site? How weathering degrees were defined? Even for a qualitative method, more detailed descriptors should be provided.

Response:

Tables 1 and 2 are the exact transcription of the LPC method (Effendiantz *et al*, 2004), coming from a report written in French. The criteria used to evaluate the *occurrence probability* and the *temporal probability* of the LPC method (Tables 1 and 2) are listed in the paper, in part 2.3.2. The criteria used are also summarized in Figure 3. In the paper, an analysis of the parameters of the LPC and SMR-based methods (*occurrence probability*, *temporal probability*, *predisposition to the failure*, volume) is proposed. The objective of this analysis is to compare the results obtained by the different groups of people. Another analysis (Principal Component Analysis, PCA) has been realized to compare the criteria used. However, the analysis did not produce any concluding results. A comment is added to discuss this point in part 3.4.

Concerning the geomorphic indicators of the SMR-based method, the evaluated level of activity can be considered as the history of the site (El-Shayeb, 1999). It is with this parameter that the *temporal probability* is taken into account. The geomorphic parameters considered correspond to all the visual indicators that provide insight on into the *temporal probability*. That can be: open fractures, signs of failure, crushed vegetation, etc….. This complementary explanation is added to the paper in the part 2.3.3 –" presentation of the SMR-based method". On the test site, the morphological traces are different, depending on the sector considered. However, in the three sectors, traces of recent failures are notable at the top of the wall.

4) Morphological traces are used by the Authors in the SMR-based method to assess temporal activity of rockfalls. However, geomorphic indicators as the spatial density of scars and presence of cracks, or other types of morphological traces merely provide data on spatial probability and, therefore, on rockfall susceptibility (e.g. Table 4 of Fell et al., 2008). Such types of geomorphic indicators do not give information on rockfall activity or temporal frequency.

Response:

In the analysis proposed in the paper, the level of activity, assessed considering the morphological traces is considered as the history of the site (El-Shayeb *et al*, 1997). Thus, it includes the temporal activity of rockfalls. Indeed, it is considered that some morphological traces can be used to evaluate the "imminent" and "very short-term" qualitative scale of the level of temporal probability (Table 2 - LPC methods). Thus, even if this approach is less global than the one proposed by the LPC, it allows the rockfall hazard level to be evaluated. We believe that the term "morphological traces" can be misleading. Thus, we propose to use the term

"weathering traces" instead of "morphological traces".

5) Weathering degree can be used to obtain relative ages of rockfalls. To be applied for estimating temporal frequency, a previous calibration by using data on temporal occurrence of events is required. Temporal data on rockfalls in the study site, or from other sites with similar setting, are also necessary for an heuristic hazard assessment to build personal judgement (tentative or fuzzy calibration). It should be clarified in the manuscript if, and how, data on recent rockfall events or on recent exposure of rock faces in the site or in close rock walls have been used for a qualitative estimation of hazard.

Response:

As presented previously, no temporal data is available on the test site. Thus, even if the estimation of temporal frequency is an interesting approach, it is not possible to apply it to the study. The objective of the study is to make the analysis in the most basic condition (e.g.: without any additional information than the one available on the test site). The scenario considered in the study presented in the paper is: an expert is in charge of the evaluation of a new site, and he does not have any temporal data available to him.

The sole data considered in our analysis is the visual evidence of events. This is discussed in the paper, in the discussion part (part 4.2 "Choice of the test site") to highlight the limits of the chosen site.

6) Susceptibility by SMR and assumed activity is used in the SMR-based method to define occurrence probability (Table 6). On one hand, this is simply formally not correct. As it has mentioned above, susceptibility (SMR) values do not give any information on probability of temporal occurrence. Nor, on the other hand, data on spatial density of morphological traces or weathering degree, unless they are calibrated with absolute ages.

Response:

As discussed previously, the vocabulary used in the paper has been modified to fit the definition proposed by Fell *et al.* (2006). Thus, using the new vocabulary associated to the proposed definition, Table 6 is formally correct. Indeed, in the proposed version of the paper, Table 6 is a combination between the SMR and the activity, and leads to evaluate a level of *predisposition to the failure*.

7) Finally, the two methods used involve relative techniques for assessment of susceptibility and qualitative temporal probability, each one having your own scales for rockfall volume and temporal probability. Though results obtained are re-scaled using a common hazard rating, they can hardly be compared if a common reference regarded as "truth" is not used. Indeed, comparison seems more reliable if is done separately for the susceptibility results and temporal probability results obtained by each technique. Comparison of susceptibility results could be carried out by using the

spatial distribution of rockfall scars (this method has been applied to validate relative susceptibility methods based on rock slope characterisation; e.g. SMR in Corominas and Mavrouli, 2009). Methods for qualitative assessment of temporal probability should necessarily be validated and compared by means of real data on temporal frequency.

Response:

The objective of the paper is to compare the level of rockfall hazard considering different assessment methods. The fact that the rockfall volume and the temporal probability scales are different is related to the methods themselves. In other words, using different assessment methods means using different scales of intensity (volume) and *temporal probability*. However, the final levels (*very low*, *low*, *moderate*, *high*) correspond to classical rockfall hazard levels (Copons et al, 2008; Bauer, 2011; Office fédéral de l'aménagement du territoire OFAT, 1997). These levels are commonly used to build hazard maps, for risk management in urban areas. The objective of the study is thus to consider different methods, and evaluate the differences in terms of levels of rockfall hazard. Moreover, the influence of the level of expertise on the result, and so on the obtained hazard levels, is also investigated. The objective of the paper is not to evaluate the "true" level of rockfall hazard, but to compare the evaluation process, considering different levels of expertise and different methods.

Based on the modifications proposed, we hope that you will be agreed to let us submit a revised version of the manuscript. We truly think that this revised version is clearer, and highlights the relevance of the study.

We will be looking forward to be hearing about your decision.

Best Regards,

The Authors

References

Bauer, M. And Newmann, P.: A guide to processing rockfall hazard from field data, ISGSR 2011, Vogt, Schuppener, Straub & Brauss (eds), 2011.

Copons, R. and Vilaplana, J. M.: Rockfall susceptibility zoning at a large scale: From geomorphological inventory to preliminary land use planning, Eng. Geol., 102(3-4), 142–151, 2008.

El-Shayeb, Y, Verdel, T. And Didier, C.: Fuzzy Reasoning for the analysis of risk in geotechnical engineering. Application to a French case. Proceedings of the

1997 International Symposium on Intelligent Design in Engineering Applications (IDEA'97). September 11, Ed. Elite Foundation, Aachen, Germany, pp 101-105, 1997.

Effendiantz, L., Guillemin, P., Rochet, L., Pauly, C. and Payany, M.: Les études spécifiques d'aléa lié aux éboulements rocheux, Laboratoire central des ponts et chaussées, Paris., 2004.

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., Savage, W. Z. and others: Guidelines for landslide susceptibility, hazard and risk zoning for land use planning, Eng. Geol., 102(3-4), 85–98, 2008.

Hantz, D., Vengeon, J. M. and Dussauge-Peisser, C.: An historical, geomechanical and probabilistic approach to rock-fall hazard assessment, Nat. Hazards Earth Syst. Sci., 3(6), 693–701, 2003.

Office fédéral de l'aménagement du territoire OFAT: Prise en compte des dangers dus aux mouvements de terrain dans le cadre des activités de l'aménagement du territoire, 1997.