Comments to the authors of the manuscript under discussion entitled "***Understanding rockfalls along the national road G318 in China: from source area identification to hazard probability simulation"***

Key finding: The presented manuscript deals a comprehensive assessment for rockfall hazard along the national road G318 in west Hubei. It combines geological surveys, disposition analysis and trajectography with Flow-R as input for rockfall hazard probability calculations.

Main concerns: The paper is a mash-up of many techniques in rockfall hazard classification/risk assessments. It starts out with rather simple and straightforward slope angle threshold techniques (executed incorrectly), adds statistical models/ such as minimum description length (MDL), multivariate logistic regression model (MLRM) and the random forest model (RFM). The flawed SAT method unfortunately leaves the comparison with more elaborate models useless. Although MLRM and RFM models may serve as additional tools, the presentation in this publication is not stringent and the sole idea of mashing all methods together does not provide a better result.

Honouring Occam's razor, if the authors want to pursue all of the mentioned techniques, they need to verify for each model addition, why it is performing better than a simpler methodology. As they fail to perform the first task of the SAT threshold analysis in a correct way, the reasoning for more complex models are insubstantial.

The calculation of rockfall hazard probability where size scenarios are independent of return periods shows a fundamental misunderstanding in combining probabilistic methods.

Combined with the excessive incorrect placement of questionable, old and mis- referenced sources, the overall manuscript – although it is based on a solid field examination and an intrinsically valuable data set – fails to deliver scientific thoroughness and quality sought for in this journal.

The authors are urged to re-assess their tool sets, assess performed analysis with great care and fit their work in a thourough background. Name dropping and big data approaches with small data sets and inconsistent analysis and error checks does not help the scientific community (and is a waste of the reviewers valuable time).

After all, I cannot recommend this article for publication. A resubmission would require substantial re-thinking of the applied methods, thorough cross-comparison of the individual tools and in-depth validation for each complexity introduced.

In the following there are some minor/technical/content corrections: (aborted after too many inconsistencies in the conception/execution/writing made a recommendation for publication impossible)

Technical corrections:

General comment on the use of parentheses: Clearly a matter of writing style, however, IMHO the excessive use of parentheses hinders the reading flow. Personal guidance is: if it's important, rephrase it into the written sentences, if it does not merit being included in the text, remove it. The authors might check their use of parentheses with this in mind, or discard it as the referee's spleen. Does not hold for introduction of acronyms, of course.

Figure font sizes: Revise the font size and general sizing of heavily loaded figures.

Abstract:

l14: kinemics → kinematics, but is it really?
l20: results agree with measurements, fit well the acquired field data, etc., but they don't show fitness
l21: size scenarios usually are linked **to recurrence** periods. The de-coupling from size scenarios to recurrence period does not make sense.

1. **Introduction**

l26ff: what about debris flows, avalanches, shallow landslides, etc?
l29: at the border between China and Nepal.
l30: crosses/leads trhough mountainous areas.
l33: book cited incorrectly, plus: does it really make sense to cite a book for common knowledge such as "rockfalls usually occur in montaineous regions"?
L38: derived from a digitital elevation model
l41: This is not the only reason for LiDAR scanning and the Fanos et al. source is clearly focused on something different (machine learning for rockfall trajectory propagation modelling)
l42: The conclusion from the cited work is rather, that it is no unambiguous SAT derivation possible. That terrain is an important basis, is common knowledge. Not many rockfalls occur in the planes.
L44: rockfall susceptibility is a combination of all of those factors. It should not be opposing, but complementary assessments.
L47: sentence makes no sense. Source areas can be identified more accurately either by using empirical, statistical or deterministic methods.
l49: how widely used is RHRS? And how accurate/universal are the proposed exponential function within the original RHRS publication? It is a method amongst many.
L51: Oommen at al. (1984) should be Bouali et al (2019)

L54-56: arguing with academic references from roughly 30 years ago, that a method is commonly used is a bit far fetched. The problematic on input data is already discussed there.
L62 ff: What is 3D collapse motion? The argument, that those models require extensive field investigation and experimental parameters as opposed to Flow-R is not substantial.

The reference Jabodeyoff et al.2003 is a link to where no manual for FLOW-R is found anymore (CONEFALL and others are found there). The statement, that FLOW-R produces more realistic results with the citations of a wrong manual is a bold – if not scientifically fraudulent - claim.
L 72: What is a fragment in this case? Usually fragments are fragmented parts from a initially released rock from the release area. Of course, those rocks are also fragments from the original rock wall etc, but in the literature, fragmentation means the breaking up of a single block during its trajectory. The influence on deposition patterns etc. is a hot topic and controversially debated. Reach angle analysis, however, can not contribute, to this discussion.
L 80ff: What is the temporal probability? Recurrence periods? There is a great many work around scenario building in rockfall etc. The authors have a point, that a thorough link between occurrence probability and scenario probability might be a weak point of current hazard mitigation literature. Please rephrase.
L85-90: Please refine the English.

2. Study Area

L93 Intence → Intense erosion and weathering processes
L94 600 m
Figure 1: all anticlines in the figure are labelled incorrectly → antivline
L101 The lithology in the area consists mainly of purplish-red mudstone
L105 Anticline → anticline
l104 frastures → infrastructure. Are there any statistics on the events on this road and the caused dammage?
L105 Anticline → anticline
L107: how obvious? What do you want to tell the reader?
L111: nucleus → core, near-wings?
L112: what is differential weathering?
L117: Figure 2a shows no vehicle damage, that is Figure 2d. “The sandstone cliff collapsed” are the steep section without vegetation, not necessarily a collapse already.

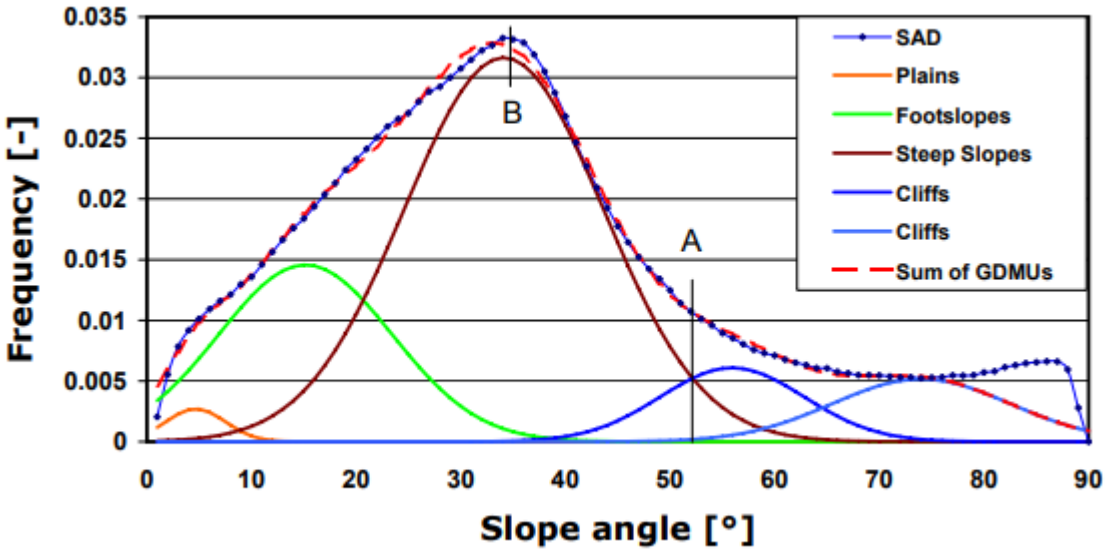
3. Methodology

Figure 3 – the presented methodology is a quite intricate interplay. A priori MLRM and RFM models work only for large data sets.

L144: SAT methodology according to Loye et al. shows quite a bit of DEM resolution dependency. The adaption of this procedure to a 10 m DEM is questionable.

L201: Reference should be Bak et al. (1988). Additionally, the reference deals with "Self-organized criticality"

L203 DOI of source Pelletier et al. 1997 is invalid.



Additionally, 27° is the transitions between footslopes and steep slopes. It is a rather low value in general as threshold.

l257: In general, the Varnes et al. (1984) citation is very old, hard to retrieve and in the context of MDLP highly likely the wrong citation.

L295: Comparison with SAT model approach is not valid, as SAT model approach is done incorrectly.

4.2 Temporal and size probability of rockfall sources

5. Discussion

l408ff: Are you altering Flow-R in order to incorporate all the promised things