Answer to EC:

We earnestly appreciate your time in reviewing the manuscript as well as your valuable comments. Please find our corrections and responses to your comments and suggestions. The corrections are listed in this response and shown in the revised manuscript.

Comments:

- **I think that the quality of the paper should be strongly improved and that several scientific steps are only inferred and not described accurately. The referee #1 pointed out: "The deformation and failure mechanisms described in the manuscript are only inferred by the authors based on field observations, but these mechanisms are not corroborated by in-situ evidences, monitoring or, at least, numerical modelling, which could be potentially highly useful to such specific purposes."**

Answer:
In previous manuscripts, the failure mechanisms was not clearly described. The stress distribution on the contact surface is difficult to be observed in reality, and there is currently no effective monitoring data. We especially thank reviewer 1 for his suggestion to carry out the related comparison work by numerical simulation. For the stress distribution, the results of numerical simulation through FLAC3D are consistent with the results of the proposed analytical method, which verifies the validity and rationality of the analytical method. A new Section has been added to the article, entitled " 4 Validation of analytical methods by numerical simulation".

The damage mechanisms at the base of the rock block play an important role in the rockfall evolution process. However, the stress distribution on the contact surface calculated by the proposed analytical methods is difficult to be validated by the field data. Therefore, numerical simulation of a biased rockfall was conducted in this study to determine the stress distribution on the contact surface between overhanging sandstone and underlying mudstone. By comparing the results of the proposed analytical methods with those obtained from the numerical simulation, the reliability of the analytical methods can be validated. FLAC3D, a professional software that utilizes the finite difference method (FDM) for three-dimensional analysis of rocks, soils, and other materials, was employed for the 3D numerical simulation. Based on the geological models, a 3D numerical simulation model was conducted with FLAC3D 6.00 to analyse the stress distribution on the contact surface (Fig. 10).

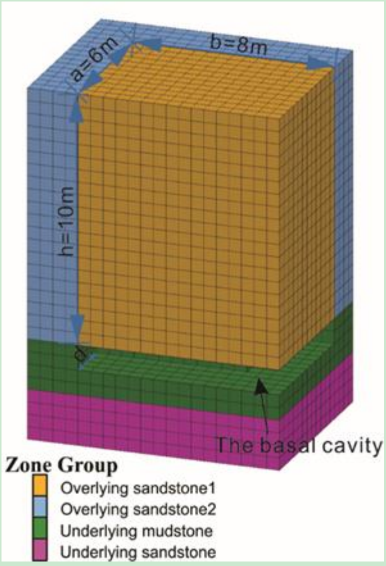


Figure 10 Numerical model built in FLAC3D

The model is mainly composed of sandstone and mudstone, which the Overhanging sandstone1 represents a unstable rock block (dimensions a×b×h are 6m, 8m, 10m respectively) , and the weathering process of the cavity is represented by excavating in stages in the underlying mudstone. Sandstone was considered as elastic model, and mudstone was assigned Mohr-Coulomb model. Material properties were determined by referring to published literature and investigation reports in the study area. The unit weight of the sandstone block (γ_s) is 25 kN/m3 (Tang et al., 2010), and the mudstone is 22.54 kN/m3. The friction angle of the contact surface (φ) is set to 25° and the cohesion (c) is set to 70 kPa (Zhang et al., 2016). Because of the strength degradation of mudstone foundations due to intense weathering, the maximum compressive stress of mudstone (σ_{cmax}) is replaced by the bearing capacity of mudstone foundations (2300 kPa), which is obtained through plate load tests in adjacent areas (Zheng et al., 2021). In addition, the maximum tensile stress of mudstone (σ_{tmax}) is valued as one-ninth of σ_{cmax}. The west, north and bottom boundaries of the model are constrained by roller boundary conditions. The cohesion and internal friction angle of the interface between Overhanging sandstone1 and Overhanging sandstone2 are set to 0. After reaching the initial force-equilibrium state, the mudstone was excavated to simulate the weathering process, and the vertical stress distribution on the sand-mudstone interface at different cavity depths was obtained, as shown in Figure 11.

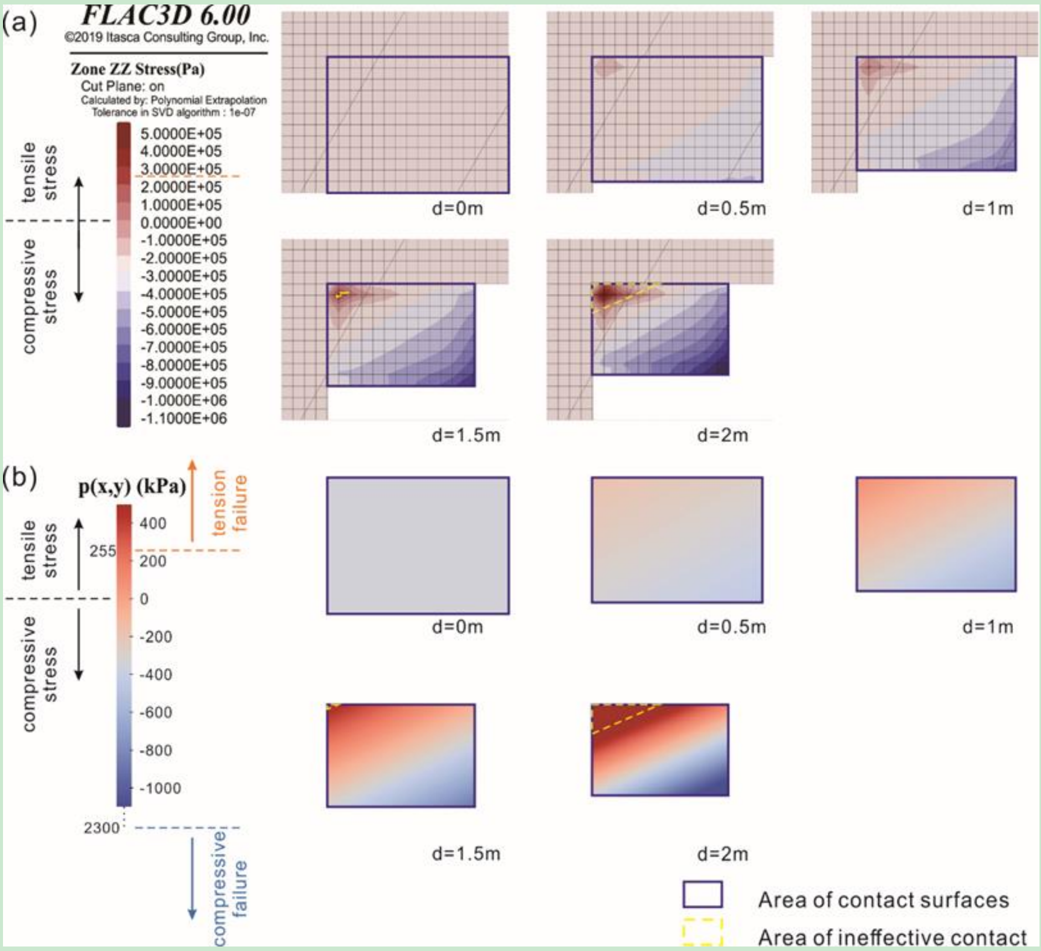


Fig.11 Diagram of stress distribution in the vertical direction on the contact interface through different methods, (a) the results of numerical simulation by FLAC3D, (b) the results of of proposed analytical method.

When there is no cavity present, represented by d=0m, the stress distribution is uniform compressive stress (According to the FLAC3D software, compressive stresses are negative). At d=0.5m, the stress remains entirely compressive, but non-uniform stress distribution occurs on the contact surfaces. At d=1m, the vertical stress value in the upper left corner of the contact interface surpasses 0 (Fig.11), indicating the presence of tensile stress. As d increases to 1.5m or 2m, the tensile stress in the upper left corner gradually intensifies, exacerbating the non-uniform stress distribution. The results obtained from the numerical simulation align with those from the analytical method, confirming the existence of tensile stress at the contact interface in the biased rockfall due to external erosion development (Fig.11). Tensile stress commonly emerges within the contact surface, making it challenging to observe directly in the field.

- **The referee #1 also pointed out an important shortcoming "Since the damage mechanisms at the base of the rock blocks play an important role in the geological context described, according to the authors, they should try to demonstrate the same mechanisms at some extent. The relationship existing between damage in the underlying mudstone and the block stability in terms of**

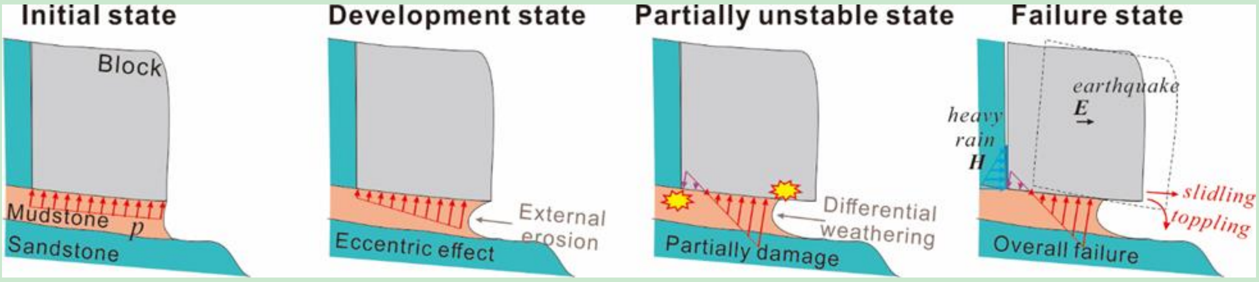
toppling and sliding mechanisms is only inferred, while the consequences are taken into account in the analytical model proposed. Such relationship should be investigated more in detail."

Answer:
The concept of ineffective contact surface is proposed in Fig. 11. With the development of differential weathering and non-uniform distributed stress, the underlying mudstone is damaged due to tensile and compressive stress in excess of strength, and the area of ineffective contact surface increases. The area that can provide anti-slip force and overturning moment is reduced, the overlying blocks are more prone to failure by sliding or toppling. The relevant content is expressed in Section.4.

In the context of the limit equilibrium method, the contact area plays a vital role in stability analysis, as shown in Eq. (21)-(30) in Section 3. The numerical simulation process provides an intuitive understanding of the influence of non-uniform stress distribution on the contact surfaces on the stability of rock blocks. Whether subjected to tension or compression, the rock layer has an ultimate strength. In Fig.11, when d=1.5m or 2m, the tensile stress exceeds the ultimate tensile strength, leading to tensile failure in the upper left corner of the stress distribution diagram. The region enclosed by a yellow dotted line represents ineffective contact, where no anti-slip force or overturning moment can be generated due to tension failure at the contact surface. Therefore, this area needs to be subtracted from the total contact area when calculating $\mathrm{Fos}_{\mathrm{sl}}$ and $\mathrm{Fos}_{\mathrm{to}}$. Similar situations occur when the compressive stress exceeds the ultimate compressive strength. The current maximum compressive stress has not reached the ultimate compressive strength in Figure 11. However, As d continues to increase, the area of compression failure will appear in the lower right corner of diagram in Figure 11. This occurrence diminishes the area capable of providing anti-slip force or overturning moment, thereby reducing the stability of the rock blocks.
The traditional LEM method does not account for distributed forces and fails to consider changes in the contact surface. The method proposed in this study addresses this issue and is applied to the calculation of the $\mathrm{Fos}_{\mathrm{sl}}$ and $\mathrm{Fos}_{\mathrm{to}}$ as presented in Eq. (21), (25) and (26)).

- **The referee #1**
  Figure 7: cavity is not a correct word to be used in the description of the specific geological situation (cavity is a void "inside" a rock mass); external erosion should work better; the overlying block should be called "overhanging".

Answer:
The content of Figure 7 has been modified. The overlying block has been changed to overhanging block in the manuscript.



- **The referee #1**
  Line 272: "according to the principle of friction" is not a rigorous expression; Eq. 21 represents the mathematical formulation of the Mohr-Coulomb criterion.

Answer:
According to the principle of friction, the ultimate shear strength →
According to **the Mohr-Coulomb criterion**, the ultimate shear strength

In addition, the abstract, introduction, conclusion and other parts of the article have been revised. Please refer to the revised manuscript for related information.