Brief Communication: Weak correlation between building damage and loss of life from landslides

Maximillian Van Wyk de Vries[1,2,3], Alexandre Dunant[4], Amy L. Johnson[5], Erin L. Harvey[4], Sihan Li[6], Katherine Arrell[7], Jeevan Baniya[8], Dipak Basnet[8], Gopi K. Basyal[9], Nyima Dorjee Bhotia[8], Simon J. Dadson[3], Alexander L. Densmore[4], Tek Bahadur Dong[8], Mark E. Kincey[10], Katie Oven[7], Anuradha Puri[8], and Nick J. Rosser[4]

[1]Department of Geography, University of Cambridge, Cambridge CB2 3EL, UK.
[2]Department of Earth Sciences, University of Cambridge, Cambridge CB3 0EZ, UK.
[3]School of Geography and the Environment, University of Oxford, Oxford OX1 3QY, UK.
[4]Department of Geography, Durham University, Lower Mountjoy, South Rd, Durham DH13LE, UK.
[5]Department of Government and Sociology, Georgia College & State University, Milledgeville 31061, US.
[6]Department of Geography, University of Sheffield, Winter St, Sheffield S37ND, UK.
[7]Department of Geography and Environmental Sciences, Northumbria University, City Campus, Newcastle upon Tyne, NE1 7RU, UK.
[8]Social Science Baha, 345 Ramchandra Marg, Battisputali, Kathmandu, Nepal.
[9]National Society for Earthquake Technology, Sainbu Bhainsepati Residential Area, Lalitpur 13775, Nepal.
[10]School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne NE17RX, UK.

Correspondence: M. Van Wyk de Vries (msv27@cam.ac.uk)

Abstract. Mapping exposure to landslides is necessary to mitigate risk and ~~reduce vulnerability~~increase resilience. Exposure maps can be constructed from building databases, akin to seismic risk assessments, but there has been little investigation of the predictive relationship between building damage from landslides and risk to human life~~from landslides~~. Our study investigates this relationship globally and in Nepal (47,213 and 5,664 landslides, respectively). While a correlation exists for nationwide totals (R^2=0.75), it is ~~near zero~~ negligible for individual events (R^2=0.025). It is important ~~not to use building datasets in isolation for~~ to not construct landslide exposure maps ~~and disaster planning to avoid unintentionally prioritising building damage~~ from building datasets alone, else building damage may be inadvertently prioritised over human lives in disaster planning.

1 Introduction

1.1 Landslides and landslide risk

Landslides cause thousands of deaths each year across a wide range of geographic environments (Petley, 2012; Kennedy et al., 2015), and are preconditioned and triggered by a wide range of anthropogenic (e.g. road cutting) and physical (e.g. earthquakes and intense rainfall) processes (e.g. van Westen et al., 2006). Landslide ~~disaster~~ risk reduction is therefore a challenging task, as it requires an understanding of a diverse range of predisposing factors, failure processes, and potential impacts over large spatial areas.

~~The~~ To date, the majority of studies ~~to date~~ focus on landslide hazard or susceptibility, constructed based ~~either~~ on a statistical analysis of past landslide records (Calcaterra et al., 2003), on datasets describing typical landslide predisposing factors (van Westen et al., 2006; Reichenbach et al., 2018), or ~~some~~ a combination of these two methods. ~~A smaller number of~~ Other studies have moved beyond this to evaluate landslide risk, combining hazard, vulnerability, and exposure (Lateltin et al., 2005; Cruden, 2018; Emberson et al., 2020). Vulnerability is the capacity to prevent, mitigate, or recover following a landslide ~~(van Westen et al., 2006)~~(van Westen et al., 2006; Alexander, 1986; Chiocchio et al., 1997; Iovine and Parise, 2002; Chen et al., 2020) , and it is commonly set at a constant value calibrated based on historical damage data or human development indicators (Atkinson, 2012). Exposure represents the spatial distribution of at-risk people, buildings, or other infrastructure.

It is instructive to compare the ~~assessment of risk from landslides to~~ landslide risk assessment workflows to those from other hazards. ~~Seismic~~ For example, the exposure and vulnerability components of seismic risk maps are commonly ~~constructed from datasets of buildings, using fragility curves to assess the vulnerability of different building types~~ drawn from building datasets (e.g., Coburn et al., 1992). This approach is justified as most earthquake deaths are directly associated with building collapse, and is appealing as it can be upscaled to large areas (Coburn et al., 1992; Doocy et al., 2013). Open-source building databases, such as OpenStreetMap, have recently ~~become available and~~ improved the availability and accessibility of this information. Various studies have considered whether seismic hazard assessments may be directly applied to landslides, or whether an analogous approach may enable these building maps to be used as an input for landslide risk models (e.g., Pollock and Wartman, 2020; Jakob et al., 2012). In the case of catastrophe risk models applied by private sector insurance companies, exposure is commonly composed of an infrastructure map with a specific economic value associated with each property (Sterlacchini et al., 2007; Atkinson, 2012). However, for building databases to be directly applicable as exposure layers for landslides, we must evaluate whether they adequately capture not only the damage ~~from landslides~~ to buildings but also ~~the~~ loss of life~~from landslides~~.

In this study, we investigate the relation between the reported total human loss from landslides (deaths and missing people) and reported building damage. We investigate this relation both on a global scale and through a ~~specific~~ detailed case study in Nepal. We consider whether building damage is a reliable proxy - with predictive value - for the total human loss from landslides, and what the implications of a decoupling between these two key indicators of impacts may be.

2 Methods

We use a harmonized database of disasters for 89 countries, DesInventar (~~Sistema de Inventario de Desastres /~~ Disaster Inventory System), to investigate the relation between landslide building damage and deaths (Atkinson, 2012; Yamazaki-Honda et al., 2019; Mazhin et al., 2021). ~~The time period covered by DesInventar varies between countries but is generally~~ DesInventar includes data from the mid-20th century to the present, ~~and despite some missing events and incomplete metadata~~ with degree of completeness and metadata varying widely between countries. Nevertheless, it remains one of the most spatially and temporally comprehensive global-scale disaster databases (Yamazaki-Honda et al., 2019; Mazhin et al., 2021). Commonly used indicators for damage are recorded, including the number of buildings destroyed and damaged, the number of lives lost, and the number

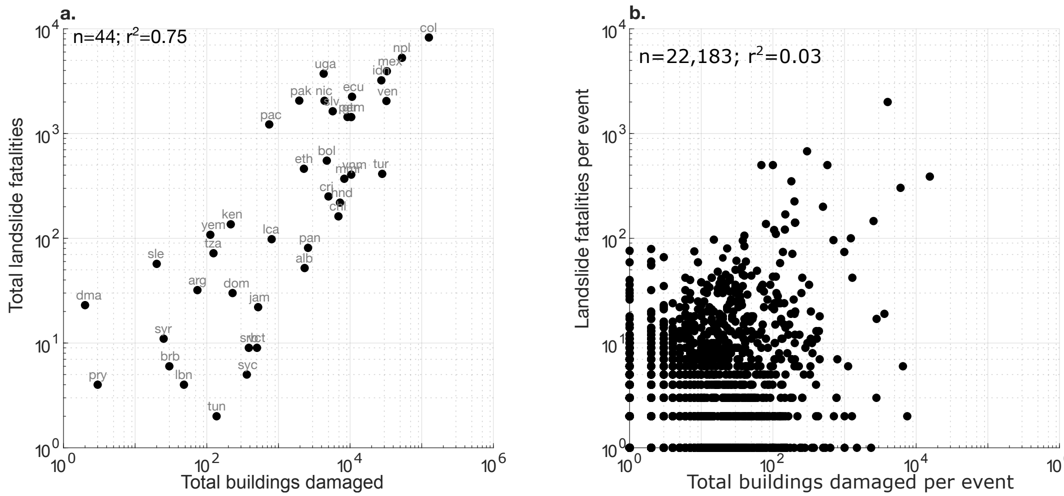

Figure 1. Plot of total human loss against total building damage for nationwide totals (a) and individual landslides (b). Country codes for in (a) are the standard nomenclatures from the DesInventar database and are included in the Supplemental Information.

declared missing for each disaster (Yamazaki-Honda et al., 2019). There is substantial work on the differing degrees of damage that landslides can do to buildings (Alexander, 1986; Chiocchio et al., 1997; Iovine and Parise, 2002; Chen et al., 2020; Del Soldato et al., , but these datasets do not allow for this level of detailed analysis and we simplify here to a binary damaged/not damaged classification. For a nationwide case study of Nepal, we supplement the DesInventar event catalogue (which ends in 2013) with comparable data from the Nepal National Disaster Risk Reduction And Management Authority (from 2014) available through the Building Information Platform Against Disaster or BIPAD Portal (NDRRMA, 2020).

We begin with a systematic search of all keywords in the DesInventar database to identify words for 'landslide' in different languages (e.g., 'deslizamiento de tierra'), spelling errors (e.g., 'landside'), or different terms for a similar physical process (e.g., 'rock slide'). A full list of the keywords used is available in the Supplemental Information. We find a total of 47,213 individual landslides, of which we exclude 25,030 which resulted in neither deaths, missing people, or building damage and include 22,183 for further analysis which record losses in at least one of those categories (deaths, missing people, and/or building damage). We regress the total human loss from each landslide (sum of deaths and number of missing persons) against the total building damage (sum of buildings destroyed and damaged) for each country with sufficient data (>10 landslides over the entire record), yielding estimates for a total of 44 countries. For Nepal, we run this workflow on both the DesInventar data alone and on a merged database comprising both DesInventar and BIPAD data.

We use different metrics to evaluate the power of ~~the predictive variable (~~total building damage ~~) towards the dependent variable (~~(predictive variable) towards total human loss (dependent variable): the coefficient of determination (R^2), the coefficient of estimation (CE), reduction of error (RE), and root mean squared error of prediction (RMSEP; Cook et al., 1994). The CE, RE, RMSEP, and relative RMSEP specifically test whether or not building damage can ~~meaningfully~~ be used as a meaningful predictor for human damage~~in predictive models~~. We adopt the formulae of Cook et al. (1994) to calculate the CE and RE, using bootstrapping to separate the data into ‘validation’ and ‘calibration’ datasets, randomly sampled (with repetition, both to the same size as the original dataset) 100 times to calculate the mean and standard deviation for the CE and RE metrics. We calculate the RMSEP and relative RMSEP from the same bootstrapped datasets. To ~~better contextualise the results for landslides~~contextualise the landslide results, we repeat the above analysis for several other disasters in the DesInventar database: floods, volcanic eruptions, earthquakes, and storms~~(shown in the Supplemental Information)~~.

3 Results

3.1 Global analysis

On a global scale, total human loss ~~and~~ positively correlates with total building damage from landslides~~are positively correlated. The goodness of fit of a linear regression line is high~~. The linear fit is good when considering the nationwide averages (R^2=0.75, F-score=112, n=44, Figure 1a), but is ~~very low when disaggregating~~ negligible when disaggregating to individual events (R^2=0.03, F-score=1250, n=22,183 Figure 1b). The F-score confirms that the relationship between the two variables is statistically significant, despite the poor correlation. This poor goodness of fit remains when considering individual landslides for each of the 44 countries with >10 events, with a median R^2 of 0.0247 (interquartile range 0.0041, 0.1017) and relative RMSEP of 590% (interquartile range 276%, 1195%). ~~The~~ For individual nations, the coefficient of estimation (median -0.22, interquartile range -4.14, -0.049) and reduction of error (median -0.049, interquartile range -0.18, 0.00) ~~calculated for individual nations~~ are also negative.

3.2 Case study: Nepal

In Nepal, an equally poor ~~goodness of fit is clear~~ fit is apparent for the combined DesInventar and BIPAD disaster inventories (Figure 2a). The regression of total human loss against total building damage in this dataset of 3,206 landslides in Nepal has an R^2 of 0.0023 and an relative RMSEP of 1279 $\pm$ 197%. The coefficient of estimation (0.0023 $\pm$ 0.0031) and reduction of error (0.0028 $\pm$ 0.0032) are ~~positive but~~ statistically indistinguishable from zero, indicating ~~little to~~ no predictive value. In simple terms, events causing the highest human loss ~~commonly~~ are not significantly associated with those which cause the most building damage, and events destroying the most buildings are not always the deadliest.

3.3 Comparison with other disasters

We repeat the same analysis with different disasters from the DesInventar database. Some, such as floods and avalanches, also have a weak relationship between total human loss and total building damage. Conversely, other disasters such as tsunamis (n = 2,126; Figure 3b) and earthquakes (n = 19,180; Figure 3c) exhibit a strong link between high total building damage and high total human loss, with the greatest human loss concentrated in the events that also affected the most properties. For lightning

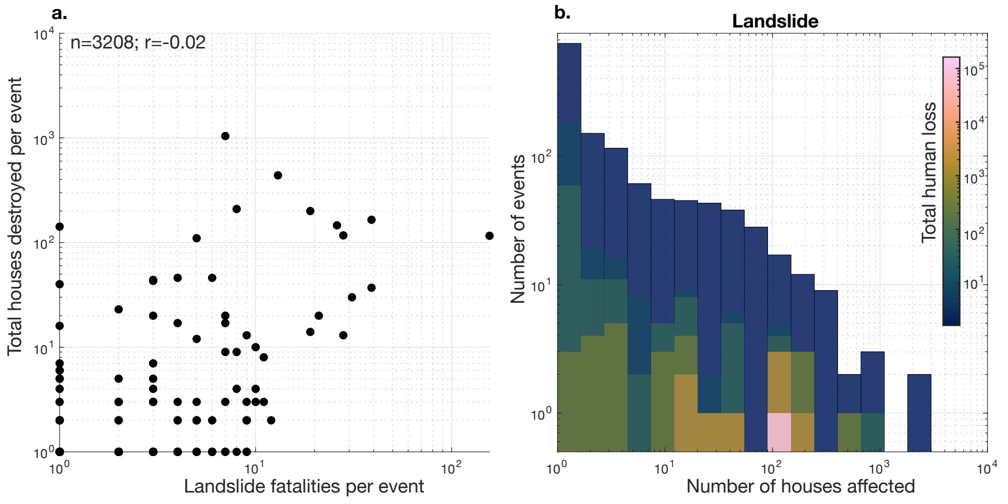

Figure 2. (a) Plot of human loss against total building damage for individual landslides in Nepal. (b) Histogram of total building damage in individual landslides for Nepal, with colours corresponding to associated total human loss. A strong correlation between building damage and human loss would result in a progressive left-to-right colour gradient and limited variation within each column. Instead, note that human loss within each column is highly variable.

strikes, we find the inverse result, with the highest total human loss in events affecting the least properties (n = 6,604, Figure 3d).

4 Discussions

~~For landslides, the~~ The correlation between human loss and ~~the~~ number of affected buildings ~~is moderate~~ from landlsides is good in national averages, but ~~poor when disaggregated into~~ negligible when disaggregated to individual events. Both human loss and total building damage follow long-tailed distributions, with most damage being accounted for by a small number of events. However, in the case of landslides, we find a poor correspondence between the events causing high levels of human loss and high levels of building damage. A comparison with other disasters shows higher correlation between these extremes for other disaster types, such as earthquakes and tsunamis~~(Supplemental figure S1)~~. Averaging the impacts of multiple events, either as temporal averages (total damage per year) or geographic averages (total damage per country) can mask this lack of correlation for individual events by cumulating both high-death, low-building damage events and low-death high-building damage events.

~~Both the databases used~~ All databases provide an imperfect and biased record of the impacts of disasters. Disasters will only be recorded if they are large or damaging enough to be noteworthy, and the size of inventories varies drastically between the 89 countries included in the DesInventar database (Yamazaki-Honda et al., 2019). Additionally, the number of deaths, missing people, and buildings damaged may be either overestimated or underestimated for events that are recorded, and the accuracy of these estimates will vary spatially and temporally (Yamazaki-Honda et al., 2019; Mazhin et al., 2021). Our study design partly mitigates these limitations, as we investigate per-disaster damage instead of the total number of events, or total damage. Even though the data sets are incomplete, we can still make inferences about the relation between total human loss and total building damage from the events that were recorded. A comparison between landslides and other disasters provides a further test, as the relation between building damage and human loss varies for different disasters. This relation is expected to be strong for earthquakes, where a large proportion of human loss is directly caused by building collapse (Coburn et al., 1992; Doocy et al., 2013), and particularly weak for lightning where buildings may shield inhabitants from damage. Both of these expectations are supported by the DesInventar database (Supplemental Figure S1).

~~The~~ Our finding that damage to properties and human loss are poorly related is consistent with the complex and geographically dispersed nature of landslides and our current understanding of their links to human mortality. Different impacts may indeed be expected depending on landslide type: for instance, large landslides with clear warning signs may damage many buildings without loss of life, while small rockfalls may cause many fatalities without damaging any buildings. Similarly, landslide early warning systems may enable effective evacuations in some parts of the world, preventing fatalities but not building damage. The exact causes of this discrepancy, including different landslide processes, effective mitigation strategies, and spatially concentrated exposure and vulnerability, are likely to vary widely across the world and within this dataset. Pollock and Wartman (2020) showed the importance of demographic, situational, and particularly behavioural factors in determining landslide morbidity, arguing that the relationship between building damage and morbidity is therefore complex. A low correlation has previously been noted between landslide-related deaths and the economic cost of landslides (Hilker et al., 2009; Kennedy et al., 2015), although this study was limited to Switzerland alone. Creating an accurate risk map relies on a combination of two components: a map of the spatial distribution of the hazard, and a measure of the exposure to this hazard. This exposure layer will depend on the objective of the risk map; for instance, insurance disaster risk maps will often be based on infrastructure value maps. In the case of landslides, our results show that building maps or databases – for example, MSBuildings or OpenStreetMap – are an inadequate proxy for the total human loss from landslides, and should not be relied upon solely to estimate risk to human lives from landsliding.

Our results show that landslide disaster mitigation strategies using risk maps constructed from infrastructure assets or building datasets may implicitly prioritise monitoring or mitigation of high-building damage events instead of high total human loss events. This raises ethical and practical issues and is generally at odds with the primary objective of disaster risk reduction programs. Two different and complementary approaches stand out to improve our understanding and representation of human loss from landsliding. The first involves building a detailed understanding of local conditions through consultations and interviews, and the second involves large-scale (and ideally dynamic) population or exposure modelling. Both methods fall outside ~~of~~ the traditional remit of landslide science, and highlight the need for ~~cross-disciplinary~~ transdisciplinary collaboration for effective landslide risk reduction.

Improving our understanding of disaster risk involves examining local risk perceptions, daily routine variations, and mitigation strategies. Local risk perceptions and actions may explain the low correlation between human casualties and building damage, either due to effective evacuation strategies reducing exposure or inadvertent actions increasing exposure (Pollock and Wartman, 2020). Additionally, considering local perspectives may reveal effective risk-reduction strategies and increase community involvement in mitigation efforts. The factors contributing to the low correlation between human casualties and the number of affected buildings are likely to vary across the 89 represented countries in the DesInventar database. Conducting informal and semi-structured interviews with local residents could help shed light on why building damage might exceed fatalities in certain landslides and vice versa within specific regions.

On a larger scale, population density and dynamic exposure maps offer an alternative perspective, albeit with some limitations. These population density maps often have low spatial resolution or rely on building data for interpolation (Lloyd et al., 2017). Additionally, static population density maps fail to capture substantial transient changes in population density occurring on a daily, seasonal, or interannual basis. In scenarios such as landslide risk assessment, where evacuation may be impractical and vulnerability is high (Kennedy et al., 2015), dynamic exposure becomes a critical element of risk-to-life modelling. Innovative modelling approaches, such as agent-based models (e.g. Zayn et al., 2020), have the potential to account for spatiotemporal population movements across different timescales. However, these methods are relatively untested in the context of disaster risk reduction, presenting an open science challenge in need of further development.

5 Conclusions

This study ~~has revealed~~ shows a complex relationship between building damage and human loss ~~in~~ from landslides. We find that, despite moderate correlation for national averages, the two are ~~poorly correlated~~ uncorrelated in individual events at both a global scale and in Nepal specifically. Therefore, building damage from landslides is not an effective predictor of the number of fatalities from the same event. Comparative analysis with other disasters highlights the contrasts between them, with a stronger link between building damage and human loss present for earthquakes and tsunamis, but not for other geohazards such as floods, avalanches, or lightning strikes. There is a need to develop exposure layers beyond simple building databases, encompassing localized insights into risk factors and dynamic population models, to improve mitigation of the deadliest landslides.

Author contributions. MV conceived the study and conducted the analyses with input from AD, AJ, EH, DL, and KA. All authors helped interpret the results and commented on the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. This research was supported by a grant from the UKRI Global Challenges Research Fund (NE/T01038X/1). We thank Daniel Costantini and one anonymous reviewer for their constructive comments, along with editor Mario Parise.

References

Alexander, D.: Landslide damage to buildings, Environmental Geology and Water Sciences, 8, 147–151, https://doi.org/10.1007/BF02509902, 1986.

Atkinson, P.: State of the art in Risk Mapping, Government, UK Government Office of Science., 2012.

Calcaterra, D., Parise, M., and Palma, B.: Combining historical and geological data for the assessment of the landslide hazard: a case study from Campania, Italy, Natural Hazards and Earth System Sciences, 3, 3–16, https://doi.org/10.5194/nhess-3-3-2003, publisher: Copernicus GmbH, 2003.

Chen, Q., Chen, L., Gui, L., Yin, K., Shrestha, D. P., Du, J., and Cao, X.: Assessment of the physical vulnerability of buildings affected by slow-moving landslides, Natural Hazards and Earth System Sciences, 20, 2547–2565, https://doi.org/10.5194/nhess-20-2547-2020, publisher: Copernicus GmbH, 2020.

Chiocchio, C., Iovine, G., and Parise, M.: A proposal for surveying and classifying landslide damage to buildings in urban areas, in: Engineering geology and the environment, pp. 553–558, https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=6240631, 1997.

Coburn, A. W., Spence, R. J., and Pomonis, A.: Factors determining human casualty levels in earthquakes: mortality prediction in building collapse, in: Proceedings of the tenth world conference on earthquake engineering, vol. 10, pp. 5989–5994, Balkema Rotterdam, 1992.

Cook, E. R., Briffa, K. R., and Jones, P. D.: Spatial regression methods in dendroclimatology: A review and comparison of two techniques, International Journal of Climatology, 14, 379–402, https://doi.org/10.1002/joc.3370140404, publisher: John Wiley & Sons, Ltd, 1994.

Cruden, D.: Landslide Risk Assessment, Routledge, google-Books-ID: nFMPEAAAQBAJ, 2018.

Del Soldato, M., Solari, L., Poggi, F., Raspini, F., Tomás, R., Fanti, R., and Casagli, N.: Landslide-Induced Damage Probability Estimation Coupling InSAR and Field Survey Data by Fragility Curves, Remote Sensing, 11, 1486, https://doi.org/10.3390/rs11121486, number: 12 Publisher: Multidisciplinary Digital Publishing Institute, 2019.

Doocy, S., Daniels, A., Packer, C., Dick, A., and Kirsch, T. D.: The Human Impact of Earthquakes: a Historical Review of Events 1980-2009 and Systematic Literature Review, PLoS Currents, 5, ecurrents.dis.67bd14fe457f1db0b5433a8ee20fb833, https://doi.org/10.1371/currents.dis.67bd14fe457f1db0b5433a8ee20fb833, 2013.

Emberson, R., Kirschbaum, D., and Stanley, T.: New global characterisation of landslide exposure, Natural Hazards and Earth System Sciences, 20, 3413–3424, https://doi.org/10.5194/nhess-20-3413-2020, publisher: Copernicus GmbH, 2020.

Hilker, N., Badoux, A., and Hegg, C.: The Swiss flood and landslide damage database 1972–2007, Natural Hazards and Earth System Sciences, 9, 913–925, https://doi.org/10.5194/nhess-9-913-2009, publisher: Copernicus GmbH, 2009.

Iovine, G. and Parise, M.: Schema illustrato per la classificazione ed il rilievo dei danni da frana in aree urbane, Memorie della Società Geologica Italiana, 57, 595–603, 2002.

Jakob, M., Stein, D., and Ulmi, M.: Vulnerability of buildings to debris flow impact, Natural Hazards, 60, 241–261, https://doi.org/10.1007/s11069-011-0007-2, 2012.

Kennedy, I. T. R., Petley, D. N., Williams, R., and Murray, V.: A Systematic Review of the Health Impacts of Mass Earth Movements (Landslides), PLoS Currents, 7, ecurrents.dis.1d49e84c8bbe678b0e70cf7fc35d0b77, https://doi.org/10.1371/currents.dis.1d49e84c8bbe678b0e70cf7fc35d0b77, 2015.

Lateltin, O., Haemmig, C., Raetzo, H., and Bonnard, C.: Landslide risk management in Switzerland, Landslides, 2, 313–320, https://doi.org/10.1007/s10346-005-0018-8, 2005.

Lloyd, C. T., Sorichetta, A., and Tatem, A. J.: High resolution global gridded data for use in population studies, Scientific Data, 4, 170 001, https://doi.org/10.1038/sdata.2017.1, number: 1 Publisher: Nature Publishing Group, 2017.

Mazhin, S. A., Farrokhi, M., Noroozi, M., Roudini, J., Hosseini, S. A., Motlagh, M. E., Kolivand, P., and Khankeh, H.: Worldwide disaster loss and damage databases: A systematic review, Journal of Education and Health Promotion, 10, 329, https://doi.org/10.4103/jehp.jehp_1525_20, 2021.

Petley, D.: Global patterns of loss of life from landslides, Geology, 40, 927–930, https://doi.org/10.1130/G33217.1, 2012.

Pollock, W. and Wartman, J.: Human vulnerability to landslides, GeoHealth, 4, e2020GH000 287, https://doi.org/10.1029/2020GH000287, 2020.

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., and Guzzetti, F.: A review of statistically-based landslide susceptibility models, Earth-Science Reviews, 180, 60–91, https://doi.org/10.1016/j.earscirev.2018.03.001, 2018.

Sterlacchini, S., Frigerio, S., Giacomelli, P., and Brambilla, M.: Landslide risk analysis: a multi-disciplinary methodological approach, Natural Hazards and Earth System Sciences, 7, 657–675, https://doi.org/10.5194/nhess-7-657-2007, publisher: Copernicus GmbH, 2007.

van Westen, C., van Asch, T., and Soeters, R.: Landslide hazard and risk zonation—why is it still so difficult?, Bulletin of Engineering Geology and the Environment, 65, 167–184, https://doi.org/10.1007/s10064-005-0023-0, 2006.

Yamazaki-Honda, R., Nair, S., Touzon Calle, I., and Serje, J.: Desinventar Sendai 10.1.2 User Manual Analysis, User Guide 10.1.2, UN Office for Disaster Risk Reduction, 2019.

Zayn, A. R., Ramdani, F., and Bachtiar, F. A.: Agent-based modelling and simulation for evacuation of landslides natural disaster, Journal of Information Technology and Computer Science, 5, 194–206, https://doi.org/10.25126/jitecs.202052172, number: 2, 2020.