Simulation-based Flood Fragility and Vulnerability Analysis for Expanding Cities

Leandro Iannacone
Research Fellow, Dept. Civil, Environmental and Geomatic Engineering, University College London, UK

Vibek Manandhar
Structural Engineer, National Society of Earthquake Technology (NSET), Nepal

Januka Gyawali
Hydrologist, Practical Action, Nepal

Prayash Malla
Structural Engineer, National Society of Earthquake Technology (NSET), Nepal

Roberto Gentile
Lecturer (Assistant Professor), Institute for Risk and Disaster Reduction, University College London, UK

Maggie Creed
Lecturer (Assistant Professor), James Watt School of Engineering, University of Glasgow, UK

Ramesh Guragain
Director, National Society of Earthquake Technology (NSET), Nepal

Carmine Galasso
Professor (Full), Dept. Civil, Environmental and Geomatic Engineering, University College London, UK

ABSTRACT: Accurately quantifying flood-induced impacts on buildings and infrastructure systems is essential for risk-sensitive planning and decision-making in expanding urban regions. Flood-induced impacts are directly related to the physical components damaged due to contact with water. Conventional fragility analysis approaches for flooding do not account for the physical damage to the individual components, mostly relying on empirical methods based on historical data. However, recent studies proposed simulation-based, assembly-based fragility models that account for the damage to building components. Such fragility models require detailed inventories of vulnerable components for building archetypes of a specific region of interest. While such inventories and building portfolios exist for developed countries like the USA, they are not suitable for low- and middle-income countries, which have different building types and contents. This is especially true for rapidly expanding cities characterised by extensive informal settlements. This paper details how to adapt the available methodologies for flood vulnerability assessment to the context of formal and informal settlements of expanding cities in the global south. It also details the development of content inventories for households in these cities using field surveys. The proposed survey is deployed in various areas vulnerable to floods in Kathmandu, Nepal. Based on the survey results, each component within the household is associated with a corresponding flood capacity (resistance) distribution. These distributions are then employed in a simulation-based probabilistic framework to obtain fragility relationship and consequence models. The relevant differences between the results obtained in this study and those from previous studies are then investigated for a case-study building type. In addition, the influence of socio-economic factors (e.g.,

household income) and past flood experience (possibly resulting in various flood-risk mitigation strategies at a household level) on the resulting flood impacts is also included in the model.

1. INTRODUCTION

Flooding events have increased over the past years due to climate change (Hirabayashi et al. 2013), leading to an increase in negative socio-economic impacts on communities worldwide (Sampson et al. 2015), especially in growing cities characterised by rapid urbanization. Poorer communities in the global south are particularly exposed and vulnerable to flooding due to a combination of factors, including adverse seasonal climate, substandard/degrading physical infrastructure, and limited ability to cope with the impacts of flooding (Tanner and Mitchell 2008). Properly quantifying the impact of floods on vulnerable communities is required to develop effective risk-management and resilience-increasing strategies. Most available models to quantify flood vulnerability are purely empirical and generally based on a qualitative assessment of post-disaster damage (e.g., FEMA 2010). However, recent literature works have shifted the focus towards approaches that use proper simulation-based fragility approaches. In particular, Nofal et al. (2020) have proposed quantifying flood's impact by looking at the fragility of the individual components within the household. In Nofal et al. (2020), capacity estimates for each component are obtained and compared with the demand values associated with the flood event (i.e., water height and flood duration). The methodology relies on data obtained from content inventories for the specific households that are the subject of inquiry, namely typical households in the United States (and countries with similar socio-economic characteristics). The damage is also quantified in terms of pure asset losses, a practice that misrepresents the actual impact on poorer communities (Howell and Elliot 2018).

This paper details how the available methodology for flood vulnerability assessment must be adapted to incorporate the specificities typically associated with formal and informal settlements in expanding communities of the global south. The provided insights are used to develop a survey form tailored to the input required by the methodology, to be deployed in relevant communities, and to develop pertinent content inventories for the associated building portfolios. As a demonstration, surveys are conducted in several communities across the Kathmandu Valley, Nepal, and the obtained content inventories are used to generate loss curves for a case-study household. The relevant differences of the proposed methodology (compared with the approach available in the literature) are then highlighted.

2. METHODOLOGY

The proposed methodology for flood vulnerability assessment builds upon the formulation proposed by Nofal et al. (2020), which looks at the damage to the individual components that make a household inventory. The demand in terms of water height H_D (for the univariate model) and height and duration H_D and D_D (for the bivariate model) is compared with the capacity (or resistance) of each i-th component, expressed in terms of its depth resistance H_{c_i} and its duration resistance D_{c_i}. In the bivariate model, each component is assumed to fail whenever it is submerged ($H_D > H_{c_i}$) for longer than its duration resistance ($D_D > D_{c_i}$). In other words, the probability of failure P_{f,c_i} of the i-th component is obtained as a function of the demand $\mathbf{D} = [H_D, D_D]$ as

$$P_{f,c_i}(\mathbf{D}) = P(H_D > H_c \cap D_D > D_c) \tag{1}$$

In the univariate model, instead, P_{f,c_i} is obtained as the probability that the component is submerged by the maximum water height across the flood event (i.e., $P_{f_{c_i}}(\mathbf{D}) = P(H_D > H_c)$). The total loss $L_T(\mathbf{D})$ due to a flood with demand $\mathbf{D}$ is computed as the sum of the costs associated with each failed component. The replacement cost of the i-th component is also modelled as a random variable L_{c_i}, such that the expected value of L_T can be found as

$$\mathrm{E}[L_T] = \sum_{i=1}^{N} \mathrm{E}\left[L_{C_i}\right] P_{f,c_i} \tag{2}$$

where N is the total number of components in the household. Alternatively, statistics of L_T can be obtained by sampling the failure and associated cost of each component in a simulation-based approach (Nofal et al. 2020).

The individual aspects of the described methodology (such as the distribution of the capacity of each component) depend on the type of households and socio-economic context of interest. The following subsections detail how to model such aspects and how the proposed approach has been adapted from the available formulations to account for the specificities of settlements in expanding cities of the global south.

2.1. Capacity of the individual components

The depth resistance of each component H_{c_i} typically corresponds to the placement height of the component within the household. Such a variable is modelled as a truncated normal distribution by Nofal et al. (2020). The assumption of a truncated normal distribution might be suitable for highly standardised environments such as US households. However, formal and informal settlements in expanding global south cities tend to exhibit a higher prevalence of disorderliness than their counterparts in more wealthy nations. Furthermore, preliminary surveys conducted by the authors in the Kathmandu Valley, Nepal, have highlighted the common practice of storing items susceptible to floods (such as food storages) directly at grade (i.e., $H_{c_i} = 0$). To address these aspects, we model the depth resistance of these components with hurdle models (Cragg, 1971) rather than truncated normal distributions. In hurdle models, random variables are modelled using two parts: (i) the probability of attaining value zero, and (ii) the probability mass function (for discrete random variables) or probability density function (for continuous random variables) of the non-zero values. Using hurdle models to model the distribution of H_{c_i} provides a compact way to express the probability that the component is stored at grade along with the distribution of its placement height when not stored at grade. The parameters of the hurdle models $\boldsymbol{\Theta}$ are calibrated using maximum likelihood estimation (MLE) based on the readings for H_{c_i} collected on the field. The likelihood is proportional to the probability of the collected values conditioned on given values of $\boldsymbol{\Theta}$. If we collect M values $h_{i,m}$ $(m = 1, \ldots, M)$ to estimate H_{c_i}, we can express the likelihood as

$$L(\boldsymbol{\Theta}) \propto \prod_{m=1}^{M} P\left[H_{c_i} = h_{i,m}\right] \tag{3}$$

Eq. (3) assumes that there is no uncertainty in the collected data $h_{i,m}$. In other words, it assumes that the surveyors can exactly measure the placement height of each component in the field. To account for scenarios where the placement height of the components cannot be measured precisely, we reformulate the likelihood as follows:

$$\begin{gathered} L(\boldsymbol{\Theta}) \propto \prod_{H.C.} P\left[H_{c_i} = h_{i,m}\right] \times \\ \prod_{M.C.} P\left[h_{i,m} - \sigma_M < H_{c_i} < h_{i,m} + \sigma_M\right] \times \\ \prod_{L.C.} P\left[h_{i,m} - \sigma_L < H_{c_i} < h_{i,m} + \sigma_L\right] \end{gathered} \tag{4}$$

where $H.C.$ denotes high-confidence data (e.g., placement height precisely measured on the field), $M.C.$ denotes medium confidence data (e.g., placement height approximately measured) with associated error σ_M , and $L.C.$ denotes low confidence data (e.g., placement height assumed or extrapolated from similar measurements) with associated error σ_L. Eq. (4) borrows from similar formulations for the likelihood that include lower bound and upper bound data in the collection (Gardoni et al. 2002). The values of σ_M and σ_L

depend on several factors, such as the surveyor's expertise and the accessibility conditions of the household. In this paper, we assume $\sigma_M = 0.1h_{i,m}$ and $\sigma_L = 0.2h_{i,m}$.

2.2. Accounting for flood prevention measures at the household level

Communities in expanding cities of the global south have developed a certain level of resilience towards recurring flood events, given their historical exposure to and familiarity with these types of natural hazards. For example, the inhabitants of countries such as India, Nepal and Pakistan are annually subjected to flooding during the monsoon season, as this climatic phenomenon brings about excessive and prolonged precipitation that exceeds the capacity of the drainage and water management systems in these regions (Loo et al. 2015). As such, flood prevention measures are usually in place in most households and should be accounted for when quantifying the impact of the flood. A preliminary survey conducted by the authors highlighted that most flood prevention measures take the form of low-level barriers to prevent water from entering the household unit. To include the effect of a barrier of height H_B on the expected losses, the total losses L_T are appropriately scaled, assuming that the damage from flood events with $H_D < H_B$ will be totally prevented, i.e.

$$\widehat{L_T} = L_T \cdot P[H_B < H_D] = L_T F_B(H_D) \qquad (5)$$

where $F_B(\cdot)$ is the Cumulative Distribution Function of H_B, obtained by fitting a hurdle model to the readings collected in the field, following the procedure described in Section 2.1 for H_{c_i}. By using a hurdle model for $F_B(\cdot)$, Eq. (5) also accounts for the probability that the household does not have any flood prevention measure in place (i.e., $H_B = 0$).

2.3. Incorporating societal consequences in loss assessment

The available works in the literature quantify the losses due to a flood as the sum of the replacement costs of the physical components damaged by the water. In other words, they quantify the impact of the flood through direct economic asset losses. While direct asset losses are a good proxy for the damage and discomfort of wealthy individuals affected by a disaster, they are unsuited to characterise the effect of hazards on the poorer communities, which are heavily physical-asset poor and affected by the post-disaster dynamics of various socio-economic variables, such as access to primary services, and disruptions to their source of income (Howell and Elliot 2018). To promote more equitable approaches to risk quantification and post-disaster recovery actions, recent works have shifted to quantifying consequences based on classical welfare economics (Hallegatte et al. 2016), using well-being losses as the relevant metric. Well-being is defined as the loss of the utility of the consumption associated with each household. Using utility rather than pure consumption accounts for the differential impact between richer and poorer individuals. To incorporate such aspects into the proposed framework for flood vulnerability assessment, we also consider the income loss for each household, Δi^H, in addition to the direct economic losses L_T as

$$\Delta i^H = \Delta i^D + \Delta i^C \qquad (6)$$

where Δi^D is the reduction of income due to the failure of components that are critical to the household's livelihood (e.g., machinery or tools associated with the occupants' profession), and Δi^C is the reduction of income due to disruptions at the community/infrastructure level (i.e., losses related to the disruptions of the social network of each household; see Cremen et al. 2022). We obtain the loss Δi^D by associating an additional random variable denoted as loss of livelihood (L_{l_i}) to each component of the household (in addition to the replacement cost L_{C_i}). Such quantity can be approximated by the loss of income that the failure of such a component would cause. As such, the expected value of Δi^D can be obtained as

$$\mathrm{E}[\Delta i^D] = \sum_{i=1}^{N} \mathrm{E}[L_{l_i}] P_{f,c_i} \qquad (7)$$

On the other hand, the value of Δi^C does not directly relate to the damage to the components within the households, but must be obtained as a function of the disruptions at the community-infrastructure level. Such disruptions are quantified based on the socio-economic dynamics of the problem of interest, e.g., the number of disrupted nodes in the social network (Cremen et al. 2022) or the loss of transportation network connectivity (Silva-Lopez et al. 2022). The specific relationships between the relevant quantities and Δi^C are outside the scope of the present work but will be investigated in future works. The total reduction in consumption of the household Δc is obtained as the sum of the direct monetary losses L_T and the loss of income Δi^H (i.e., $\Delta c = L_T + \Delta i^H$). The consumption after the occurrence of the disastrous event (c^*) can then be found as

$$c^* = c_0 - \Delta c \tag{8}$$

where c_0 is the pre-disaster consumption. The consumption is then translated to utilities using the following equation:

$$u^* = \frac{{c^*}^{1-\eta}}{1-\eta} \tag{9}$$

where η describes the elasticity of the utility value of a marginal unit increase in consumption and allows to account for the larger impact of the disaster on poorer communities (Markhvida et al. 2020). Finally, the well-being loss ΔW is obtained as the difference between the pre-disaster and the post-disaster utility, i.e. $\Delta W = u_0 - u^*$. It is worth noting that this formulation of the well-being loss is alternative to the one in Markhvida et al. (2020), where ΔW is computed by integrating over the entire system recovery period. This alternative formulation provides a good proxy for the effect of the flood in terms of well-being without requiring any recovery model.

3. SURVEY DEVELOPMENT AND IMPLEMENTATION

To collect data to develop an inventory for formal and informal households in expanding cities of the global south, we designed a survey form tailored to the input required by the methodology described in the previous section. The information collected in the form is detailed in Table 1.

Table 1: Parameter values.

Data collected	**Details**
Type of settlement	Formal, semiurban, informal
Location	Address, GPS coordinate, distance from river
Previous flood	Depth, damage, type, sediment deposit, drinking water availability, water logging, electricity disruption, economic loss, immediate response
Residents	Gender distribution, age distribution, migrant/ indigenous, mobility issues
Building plan	Sketch (labelled windows, doors, raised walls, location of critical items)
Embankment/raised wall at entrance	Height
Building characteristics	No. of storeys above/below ground, category, occupancy, material, wall type, roof shape, roof type, flooring type, floor insulation, cladding, doors and windows type, staircase, structural condition
Placement height of components	Number, height, measurement confidence, estimated cost. (Categories: partitions, appliances, furniture, food storages and live animals, floor/ceiling/ roof members, wall finishing, wall opening, sanitary, miscellaneous items susceptible to flood)

To address the specificities of the formal and informal settlements in the global south, we

- collected information about the flood events previously experienced by the surveyed household. Namely, we collected information about the type of the previous flood (flash, surface, river), whether there was any sediment deposit, and whether water availability was affected. The height of barriers for flood prevention was also collected, consistently with the methods in Section 2.2.

- collected the height of each component in the household which could be damaged by the occurrence of the flood. We also noted the precision of each measurement (high, medium, low), expected replacement cost and loss of income caused by the component's failure. These values were estimated by consulting members of the local communities.
- collected information about the members of the family units to better estimate the effects of a flood in terms of socio-economic impact on household livelihood.
- identified components in the household that are essential to livelihood (e.g., machinery required to conduct regular occupational tasks, merchandise in case of surveyed shops).

The survey was conducted in January and February 2023 in 50 households in formal settlements and 50 households in informal settlements in the Kathmandu Valley, Nepal. For formal settlements, 25 households were surveyed in ward 26 of Samakhusi, and 25 households were surveyed in ward 4 of Madhyapur Thimi. The location at Madhyapur Thimi suffered heavy flooding in 2018 when the Hanumante river overflowed, while the location at Samakhusi experienced frequent flooding every other year from the Bishnumati and Samakhusi River. These locations were also considered ideal because they offered a good distribution of the intended typology of households and businesses. Similarly, for informal settlements, 25 households in ward 1 of Madhyapur Thimi along the Manohara River and ward 11 of Kathmandu Metropolitan City along the Bagmati River were surveyed. These are some of the informal setllements in the valley that get flooded every year during monsoon, and also represent heterogeneity in terms of flooding depth, exposure and vulnerability. Figure 1 shows a map of the surveyed locations within the Kathmandu Valley. The data collected in the surveys were digitised and used to generate the input required for the methodology described in Section 2.

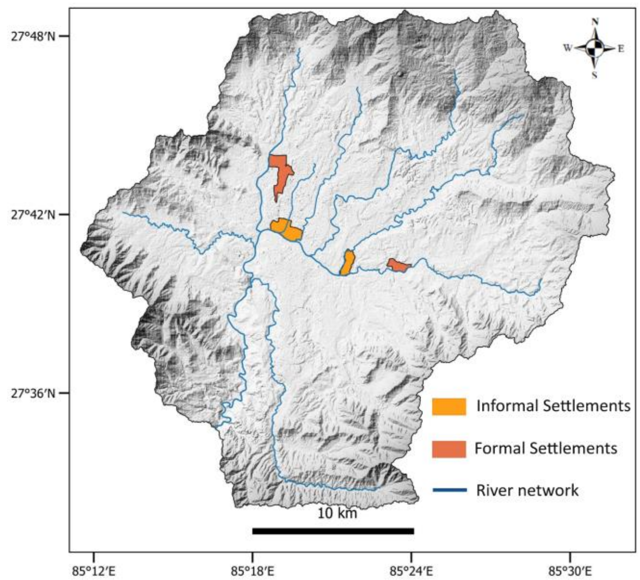

Figure 1: Survey locations in the Kathmandu valley, Nepal

4. PRELIMINARY FINDINGS

To highlight the relevant implications of the proposed methods, a simplified household is analysed under the assumptions in the flood vulnerability framework in Nofal et al. (2020) and the assumptions introduced in this paper. The following components are: fuse box, electrical switches, light fixtures, food storages, computer/laptop, beds, TV set, stove/electric cookers, refrigerator. The variables in Table 2 are considered for each component.

Figure 2 shows the loss curve in terms of asset (monetary) losses (obtained as the sum of replacement costs) for a typical household in the US (archetype 1 in Nofal et al. 2020), and a typical household from a formal settlement in Kathmandu, Nepal. Both curves were obtained under the same assumptions (truncated normal distributions for placement heights, no flood prevention measures). All costs are translated in terms of US dollars. As expected, the asset losses of a typical household in the US greatly exceed the ones for a typical household in Nepal.

Table 2: Variables associated with each parameter.

Variable	Nofal et al. (2020)	Present study
Placement height	Supplementary material: Building archetype 1	Consistent with data collected from surveys
Replacement cost	Supplementary material: Building archetype 1	Estimated by consulting the local communities
Loss of livelihood	-	Estimated by consulting the local communities

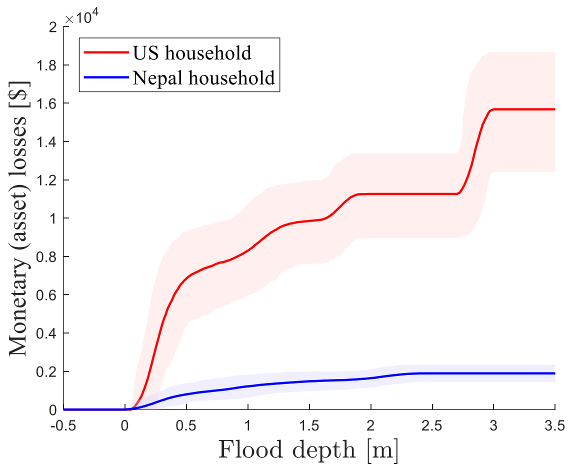

Figure 2: Asset losses of a typical US household and a typical household in Nepal

Figure 3 shows the implications of the proposed assumptions in terms of the monetary losses of the Nepali household. Monetary losses are now obtained as the sum of the replacement costs and the losses of livelihood (in terms of reduced income due to failed components). Reductions of income due to disruptions at the community-infrastructure level are neglected for this case study. The adjusted curve in Figure 3 is also obtained by modelling the placement heights of the components with hurdle models, and accounting for the possible presence of flood prevention measures (in the form of a barrier preventing water from entering the household). The adjusted curve shows how neglecting these factors could lead to an underestimation of the monetary losses. This is especially true for minor floods ($H_D < 0.5$ m), the most commonly experienced by the investigated households.

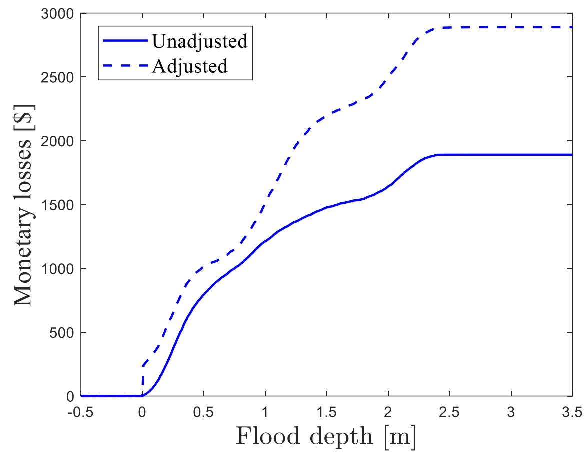

Figure 3: Effect of the proposed assumptions on the monetary losses of a household in Nepal

Finally, Figure 4 shows the expected losses in terms of well-being for a typical US household and a typical Nepali household.

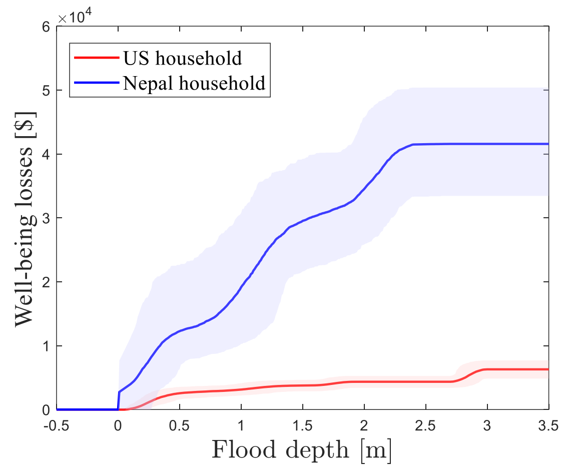

Figure 4: Well-being losses of a typical US household and a typical household in Nepal

The curve for the Nepali household is obtained from the adjusted curve in Figure 3. The initial consumptions are assumed as $c_0 = 100000$ \$ for the US household and as $c_0 = 10000$ \$ for the Nepali household. The elasticity for the utility is assumed as $\eta = 1.5$ in line with Hallegatte et al. (2016). For interpretability purposes, the well-being losses in terms of utility are converted into equivalent well-being losses in terms of consumption, ΔW_c, using the following equation (Markhvida et al. 2020):

$$\Delta W_c = \Delta W / \frac{\mathrm{d}u}{\mathrm{d}c} |_{c_{\mathrm{mean}}} \quad (10)$$

where $c_{\mathrm{mean}} = 50000$ \$.

5. CONCLUSIONS

The paper investigated the development of flood loss curves for households in expanding cities of the global south. The proposed methodology builds upon formulations available in the literature by incorporating several aspects related to the peculiar socio-economic characteristics analysed. In particular, the placement height of the components in the households was modelled with hurdle models to account for the multiple components typically stored at grade. Maximum Likelihood Estimation was used to estimate the parameters of the models, incorporating the uncertainty of the collected data. Furthermore, the chance of having flood prevention measures in place was also considered. Finally, curves were obtained in terms of well-being losses to account for the differential impact between the poorer and the richer households. A survey form was developed to collect the input required to run the proposed models, and surveys were deployed in formal and informal settlements of the Kathmandu valley, Nepal. Preliminary findings based on the results of the surveys were shown in this work, and the collected data will be used for portfolio loss analyses of vulnerable communities in expanding cities of the global south.

6. ACKNOWLEDGEMENTS

The authors acknowledge funding from UKRI GCRF under grant NE/S009000/1, Tomorrow's Cities Hub. The authors also thank surveyors, ward offices, and Mahila Ekata Samaj for helping with the collection of data.

7. REFERENCES

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, 829-844.

Cremen, G., Galasso, C., & McCloskey, J. (2022). A Simulation‐Based Framework for Earthquake Risk‐Informed and People‐Centered Decision Making on Future Urban Planning. *Earth's Future*, 10(1), e2021EF002388.

FEMA, F. (2010). Multi-hazard loss estimation methodology. *Earthquake Model/Technical Manual*.

Gardoni, P., Der Kiureghian, A., & Mosalam, K. M. (2002). Probabilistic capacity models and fragility estimates for reinforced concrete columns based on experimental observations. *Journal of Engineering Mechanics*, 128(10), 1024-1038.

Hallegatte, S., Vogt-Schilb, A., Bangalore, M., & Rozenberg, J. (2016). *Unbreakable: building the resilience of the poor in the face of natural disasters*. World Bank Publications.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., ... & Kanae, S. (2013). Global flood risk under climate change. *Nature climate change*, 3(9), 816-821.

Howell, J., & Elliott, J. R. (2018). As disaster costs rise, so does inequality. *Socius*, 4, 2378023118816795.

Loo, Y. Y., Billa, L., & Singh, A. (2015). Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia. *Geoscience Frontiers*, 6(6), 817-823.

Markhvida, M., Walsh, B., Hallegatte, S., & Baker, J. (2020). Quantification of disaster impacts through household well-being losses. *Nature Sustainability*, 3(7), 538-547.

Nofal, O. M., van de Lindt, J. W., & Do, T. Q. (2020). Multi-variate and single-variable flood fragility and loss approaches for buildings. *Reliability Engineering & System Safety*, 202, 106971.

Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., & Freer, J. E. (2015). A high‐resolution global flood hazard model. *Water resources research*, 51(9), 7358-7381.

Silva-Lopez, R., Bhattacharjee, G., Poulos, A., & Baker, J. W. (2022). Commuter welfare-based probabilistic seismic risk assessment of regional road networks. *Reliability Engineering & System Safety*, 227, 108730.

Tanner, T., & Mitchell, T. (2008). Introduction: Building the case for pro-poor adaptation. *IDS Bulletin*, 39(4), 1-5.