Designing Equitable and Dynamic Rapid Earthquake Loss Assessment Systems

Vitor Silva
Principal Investigator, Dept. of Civil Engineering, University of Aveiro, Aveiro, Portugal

Amir Taherian
Doctorate Student, Dept. of Civil Engineering, University of Aveiro, Aveiro, Portugal

Romeu Vicente
Associate Professor, Dept. of Civil Engineering, University of Aveiro, Aveiro, Portugal

ABSTRACT: After the occurrence of destructive seismic events, there is an urgent need for reliable and rapid information regarding the expected impact in terms of damage, economic losses, fatalities, and injuries. Systems for the rapid estimation of earthquake impact can be a fundamental tool in disaster risk management. However, despite the rise in the availability of low-cost sensors and recent advancements in machine learning technology, the design of near-real time loss assessment systems still follows traditional approaches, which ignores the spatial distribution of earthquake risk, and results in frameworks that tend to neglect the least vulnerable communities. In this study, considering the Metropolitan Area of Lisbon (Portugal), we investigated how ground motion sensors should be spatially distributed to minimize costs, while maximizing accuracy and reliability. To determine the optimal location of the system of sensors, we considered different risk metrics to capture the needs of distinct stakeholders. These include economic losses, number of damaged buildings per occupancy type, and number of people injured according to their socio-economic characteristics. These results can support the design of future systems, or the improvement of existing frameworks.

Seismic networks are a fundamental tool to better understand the potential consequences of seismic events and development of disaster risk management measures. Once an event occurs, these networks can support the rapid assessment of the expected losses, damage, injuries, and fatalities (e.g., Jaiswal and Wald 2013; Silva et al. 2015). Impact information shortly after the occurrence of destructive events can support local authorities in deciding where to deploy response teams, advise the healthcare system about the need to activate contingency plans, inform the (re)insurance sector about eventual large claims, and support international organizations such as the European Union or the World Bank in releasing financial support.

Several rapid loss assessment systems have been developed in the past at the urban (e.g., Istanbul - Erdik et al. 2003), regional (e.g., Romanian-Bulgarian border - Erduran et al. 2012), national (e.g., Japan - Hoshiba et al. 2011; Portugal - Silva et al. 2015; United States - Wald et al. 2020), continental (Europe - RISE project - www.rise-eu.org) and global (PAGER - Jaiswal and Wald 2010; WAPMERR - Wyss 2004) level. In general, these systems rely on three main components to estimate losses and/or building damage: 1) an exposure model characterizing the location, value, and vulnerability class of the assets; 2) a set of fragility or vulnerability functions establishing the likelihood of damage or loss conditional on an intensity measure (IM); and 3) sets of ground motion fields (also known as hazard footprints). The estimation of the ground shaking can be performed following different approaches, spanning from the simple application of a GMM using an epicenter estimated shortly after the occurrence of the seismic event (e.g., Silva et al. 2015), to a more complex procedure that incorporates data from seismic stations to reduce the bias and uncertainty in this component (e.g., Engler et al. 2022).

Amongst these three components, the estimation of the ground shaking is arguably the most influential factor. The impact of the epistemic and aleatory uncertainty from GMMs in earthquake scenarios has been thoroughly discussed in the literature (e.g., Fiorini et al. 2012, Silva 2015). These studies demonstrated that different GMMs can lead to differences in the risk results by a factor in the range of 2 to 4. On the other hand, the influence of the aleatory uncertainty can lead to a range of risk results spanning 1 order of magnitude. For example, Fiorini et al. (2012) estimated the number of unusable buildings considering the characteristics of the 2009 M6.3 L'Aquila earthquake (on soil) and indicated values ranging from 1,000 to almost 45,000 buildings. This degree of variability in the loss and damage estimates renders the results less useful for a wide range of purposes. A localized impact might be dealt by local authorities, while a more widespread impact might require assistance from the central government or even the international community. Using the example from Italy, 1,000 unusable buildings could translate into 3,000 people left homeless, while 45,000 could represent more than 100,000 people displaced.

It is thus fundamental to explore techniques to reduce the uncertainty and bias in the ground shaking component. Silva and Horspool (2019), Worden et al. (2018), and Engler et al. (2022) demonstrated how data from seismic stations can be combined with spatial and inter-period correlation models to significantly improve the reliability of ground shaking estimates in regions affected by destructive events. However, the performance of such techniques depends heavily on the density and quality of the seismic network in the affected region. Dense seismic networks only exist in a few regions in the world (e.g., Japan, Taiwan, West of the United States, Italy). Most frequently, less than half a dozen seismic stations exist in any given region. For example, Portugal has 40 stations in the continental territory (Carrilho et al. 2021), with only 7 covering the district of Lisbon where 2.1 million people live. The paucity of seismic stations is partially due to the costs of installing and maintaining a large number of devices. However, the rise in low-cost sensors capable of detecting and measuring ground shaking data in a continuous manner is changing the way in which rapid loss assessment systems can be designed. Due to their ability to transfer data in real time to a centralized system using cellular technology and the fact that they do not rely on a local power supply, denser seismic networks can be rapidly deployed in any given region at a relatively low cost.

In this study, we demonstrate the impact of incorporating data from seismic stations in the estimation of losses, using the district of Lisbon as a case study. For an earthquake scenario similar to the well-known 1755 Lisbon earthquake, we estimate the spatial distribution of economic losses, damages, and fatalities considering different densities of the seismic network, as well as different criteria for the definition of the locations of the stations. This study demonstrates how incorporating this technology can effectively improve the accuracy and reliability of loss estimates, as well as how considering specific risk metrics for the allocation of seismic stations can lead to different performances. This study is carried out within the scope of the ASSIMALTE project, funded by the Portuguese Foundation for Science and Technology, in which several seismic sensors will be deployed in the Metropolitan Area of Lisbon.

1. CONDITIONING GROUND MOTION FIELDS

The main source of uncertainty in the evaluation of earthquake scenarios is undeniably in the ground shaking component (e.g., Crowley et al. 2008; Silva 2016). The estimation of the ground shaking is characterized by two types of uncertainty: epistemic and aleatory. The former is related to our lack of knowledge or data regarding the expected ground shaking in the region, and it is usually incorporated in earthquake scenarios by considering multiple GMMs. In theory, with additional data, this source of uncertainty could be minimized. The second type of uncertainty is related to the randomness of the ground shaking process, and it cannot be completely reduced. A ground shaking intensity measure (IM) at a given site is typically estimated using the following equation:

$$Log(IM) = \mu(M, R, \theta) + \zeta \qquad (1)$$

Where IM stands for an intensity measure such as peak ground acceleration (PGA), peak ground velocity (PGV), spectral acceleration (SA), amongst others. $\mu(M, R, \theta)$ represents the mean logarithm of the estimated ground shaking, as a function of the magnitude (M), some measure of distance (R) and other parameters (θ) such as local site conditions or faulting mechanism. ζ represents the total residual of $Log(IM)$. By treating the total residual as a linear mixed-effects model (Abrahanson and Youngs 1992; Engler et al. 2022), equation (1) can be re-written as follows:

$$Log(IM) = \mu(M, R, \theta) + B + W \qquad (2)$$

Where B stands for the between-event residual (i.e., representing the variability in the ground shaking amongst events with the same characteristics) and W refers to the within-event residual (i.e., representing the variability within the same event, for sites at the same distance and soil conditions). These residuals typically follow a normal distribution, and for the assessment of earthquake damage or losses, they can be randomly sampled or numerically integrated using the respective standard deviations associated with each GMM.

Ideally, a dense network of seismic stations would exist in the regions affected by destructive earthquakes that would allow risk modelers to estimate the expected impact with little to no uncertainty in the ground shaking component. In fact, in these idealized (but certainly unrealistic) cases, GMMs would not be required. In reality, the density of seismic networks is frequently insufficient, even in countries or regions with well-developed networks such as Japan, Taiwan or west of the United States. In these cases, it is necessary to employ GMMs to estimate ground shaking in the regions far from the recording stations. However, as discussed in several studies (e.g., Worden et al. 2018; Silva and Horspool 2019; Engler et al. 2022), spatial and inter-period correlation models can be used to reduce both the bias and the uncertainty in the estimation of the ground shaking at the sites of interest. In this study, we evaluate the benefits in incorporating data from seismic stations following the procedure proposed by Engler et al. (2022). In this process, the between-event (B) term (also known as the bias) for each seismic event and IM is calculated based on the data from (hypothetical) seismic stations, while the within-event residual (W) at every location is calculated considering the data from these stations and spatial and inter-period correlation models.

2. DESIGNING RAPID LOSS ASSESSMENT SYSTEMS

As previously explained, a rapid loss assessment system requires an exposure model covering the region of interest, a set of vulnerability or fragility functions for the assets within the exposure model, and an efficient procedure to rapidly estimate the spatial distribution of ground shaking in the affected region. In this study, we use the Metropolitan Area of Lisbon to demonstrate how different criteria to design the rapid loss assessment system can affect the reliability and accuracy of the loss and damage results. We use the exposure model developed within the scope of the H2020 European project SERA (Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe - Crowley et al. 2020). This model used housing and socio-economic data from the 2011 national census survey of Portugal to create an exposure model at the smallest available administrative level ("*freguesia*"). This model was then further disaggregated on an evenly spaced grid, as illustrated in Figure 1. For the vulnerability component, we use the vulnerability functions from the global vulnerability database supported by the Global Earthquake Model (GEM - Martins and Silva 2020). These functions have been developed following an analytical approach, which uses nonlinear dynamic analysis and single-degree-of-freedom oscillators to derive fragility functions through cloud analysis.

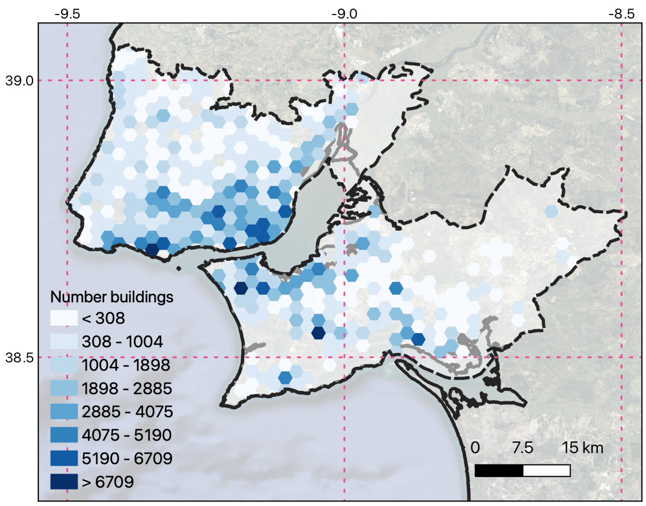

Figure 1: Spatial distribution of the number of buildings on an evenly spaced grid.

For the estimation of the magnitude and location of the event, we assumed that either the current seismic network in the region (Carrilho et al. 2021), potentially complemented by the networks proposed within this study, would be capable of estimating these parameters with a satisfactory accuracy. These calculations are already supported by the Portuguese Institute of the Sea and Atmosphere (IPMA - www.ipma.pt/pt/geofisica/sismicidade), as well as by several European and International seismological services.

For the estimation of the ground shaking in the region, ideally data from past events recorded using a dense network would be used, thus allowing the evaluation of the performance of different GMMs and the benefits of incorporating data from seismic stations. Since such data does not exist for Portugal, we generated ground shaking data using the stochastic approach proposed by Boore et al. (2003), and widely used in other countries of relatively low seismicity (e.g., United Kingdom - Rietbrock et al. (2013); Switzerland - Edwards and Fäh 2013). We considered one earthquake scenario for these analyses: an offshore strong magnitude (M8.5) event located 300 km Southwest of Lisbon. The location and magnitude of this event is identical to one of the possible solutions of the historical 1755 Lisbon earthquake, which caused widespread damage in the Southwest of the country. This method to calculate the ground shaking is hereafter termed as Approach A and will be used to benchmark the earthquake impact in this study. The spatial distribution of the (mean PGA) ground shaking for this event is illustrated in Figure 2.

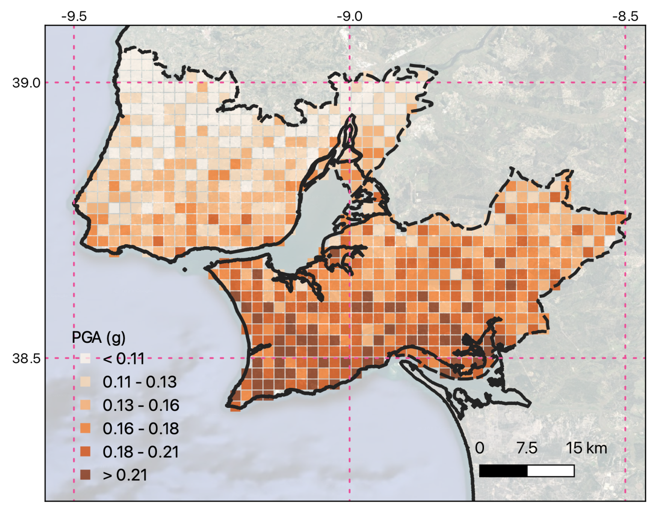

Figure 2: Spatial distribution of the ground shaking (in terms of PGA) due to the M8.5 scenario.

The generated ground shaking in the Metropolitan Area of Lisbon varies between 0.11g and 0.29g in the entire region. This relatively small range of PGA in the affected region is due to the large distance of the earthquake rupture.

Using these three components, we used GEM's OpenQuake-engine to estimate economic losses, number of damaged buildings, and fatalities. Given that for these analyses we assumed that there was no uncertainty in the ground shaking (i.e., one realization), we defined these results as the benchmark values (or the "true" impact). The results for the economic losses (in EUR) on the evenly spaced grid are presented in Figure 3.

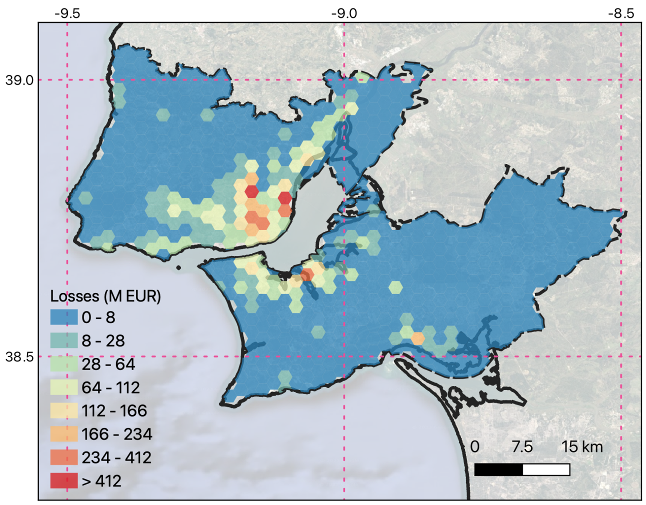

Figure 3: Spatial distribution of the economic losses due to the M8.5 offshore scenario.

To simulate the performance of a set of rapid loss assessment systems with different characteristics, we estimated the same impact under different assumptions. First, we assumed that data from seismic stations would not be considered in the estimation of the losses and damage. Instead, the system simply estimates the magnitude and location of the event, and uses a set of GMMs to calculate the spatial distribution of the ground shaking in the affected area. The ground shaking from the GMM with the lowest nominal bias (or between-event residual - B) is used for the impact assessment. Since the uncertainty from the GMM is not adjusted in this process, the resulting ground shaking is characterized by a large uncertainty (i.e., the total sigma the GMM). This method is hereafter termed as Approach B.

Then, we assumed that data from seismic stations will be incorporated in the estimation of the ground shaking, as a way to reduce the bias and uncertainty in this component. For this modeling approach, we assumed that the data recorded at the seismic stations is the ground motion records generated using the stochastic procedure, as previously described. We used the GMM with the lowest bias to estimate the ground shaking conditioned on the data from the seismic stations, following the procedure proposed by Engler et al. (2022). However, the fundamental question is how many stations should be used, and where should they be located. We follow two options to address this problem. First, we assumed an evenly spaced grid with two spatial resolutions. This approach assumes that there are no limitations in terms of the number of stations, and thus every grid cell can have a seismic station. This method is hereafter termed as Approach C. Second, we assumed that there is a limit on the number of stations that can be deployed in the region (maximum of 10), and that the location of the stations should be strategically distributed to reduce the bias in the risk metrics. The criteria to distribute stations throughout a given region might be controversial, given that some stakeholders will favor reliability in the estimation of the economic impact (e.g., insurance sectors, ministry of finance), while others will prefer a greater accuracy in the assessment of casualties (e.g., civil projection authorities). Within the evaluation of the human impact, different criteria can also be defined. For example, the reduction of the uncertainty in the estimation of human losses can focus on vulnerable groups, such as the elderly, ethnic minorities, or population with disabilities. These population groups might require special support in the aftermath of an earthquake due to mobility and/or communication difficulties. In this study, we demonstrate how the locations can change depending on whether economic or human losses are targeted. To this end, we estimated the impact for the event, and then distributed the stations proportional to the losses, while ensuring that no stations are less than 2 km apart. This strategy means that knowledge about earthquake impact is required, and thus the spatial distribution of the stations are conditional on this particular event. In reality, a probabilistic approach has to be pursued to consider all possible events in the region, as further discussed in the concluding remarks. These different rules aim at improving the reliability and accuracy of the aforementioned risk metrics. This method is hereafter termed as Approach D. A summary of the modeling approaches is presented in Table 1. We tested a number of other options to distribute stations in the region (i.e., proportional to the population density, economic value, number of buildings), but such criteria did not lead to relevant results, and are not presented here.

Table 1: Summary of the modeling approaches for the estimation of the ground shaking

Approach	*Estimation of ground shaking*	*Incorporation of stations*
A	*A stochastic method is used for the estimation of the ground shaking in the region.*	*The ground shaking is known at every location, which is equivalent to having a seismic station at every location.*
B	*One GMM is used to estimate the ground shaking in the region, based on a magnitude and epicenter.*	*No stations are considered.*
C	*The GMM with the lowest bias is used to compute the conditioned ground shaking in the region, following the procedure proposed by Engler et al. (2022)*	*Seismic stations distributed on an evenly spaced grid with two spatial resolutions: 4x4, and 8x8 km^2.*
D		*Seismic stations are distributed according to two rules: proportional to the economic losses and deaths.*

3. RESULTS

Economic losses and deaths were computed for all cases, and the results were compared to the benchmark method (Approach A). The differences in percentage (i.e., error) are presented in Figure 4.

The results indicate, in general, a small error when compared to the risk metrics computed using the "true" ground shaking (Approach A). This is partially due to the fact that a GMM with a small bias was used for these calculations. However, although not presented here, the variability in the results is relatively large for Approach B, given that the aleatory uncertainty is not constrained.

Regarding Approach C, it is interesting to note that conditioning the ground shaking on a 4x4 km^2 grid reduces the error substantially (i.e., by half and by one quarter for economic losses and deaths, respectively), while there are no significant benefits in using a network of stations with an 8x8 km^2 spatial resolution. These differences are due to two reasons: 1) when using an evenly spaced grid, more than half of the stations are located in regions with no exposure, which reduced the impact in conditioning the ground shaking; and 2) the spatial correlation between two sites decays considerably beyond 5 km. Thus, a sparse seismic network is unlikely to strongly influence the adjustment of the ground shaking for sites beyond this distance.

Finally, concerning Approach D, placing the 10 seismic stations strategically based on the expected economic losses and deaths allows effectively reducing the error, with the difference in the economic losses reducing even more than in the case of the 4x4 km^2 grid. We note that in Approach D, only 10 stations are used, while in Approach C 245 stations were considered to ensure the 4x4 km^2 spatial resolution.

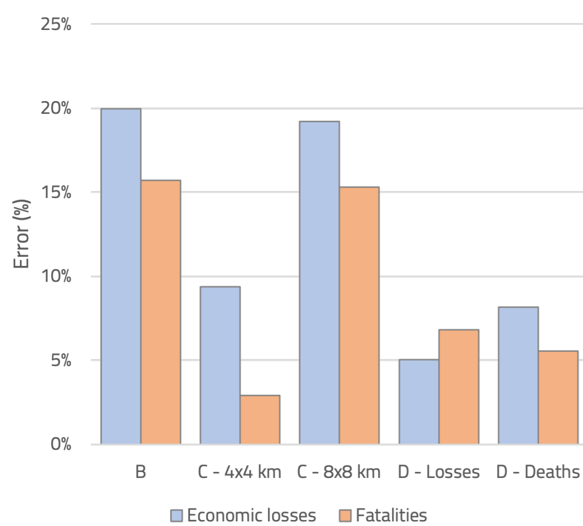

Figure 4: Differences in the impact between Approach B, C and D and the benchmark results (i.e., Approach A).

4. CONCLUSIONS

As witnessed in the recent M7.8 Turkey-Syria earthquake, reliable information regarding the expected economic losses, level of destruction, and casualties is fundamental for a multitude of tasks concerning both the public and private sector. Systems for the rapid assessment of earthquake impact are a fundamental tool in the aftermath of destructive events, as demonstrated in several countries in the past.

The results from this study highlight the potential reduction in loss estimates due to the incorporation of seismic station data in the calculation of the spatial distribution of ground shaking in the affected region. We investigated different strategies to distribute seismic stations in the region and concluded that distributing sensors in an evenly spaced grid only leads to benefits when the distance between stations is relatively short. This finding is in agreement with the correlation length indicated by most spatial correlation models. We tested other strategies that involve positioning sensors proportional to the number of buildings, economic value, population density, deaths, and economic losses. We noted that for this particular seismic scenario, despite the large region, distributing 10 stations proportional to the economic losses or fatalities led to benefits identical to those when using 245 stations on an evenly spaced grid. However, this option has the disadvantage of requiring data concerning the impact of the seismic event a priori. Nonetheless, we argue that the distribution of seismic stations in a given region can be performed using the results from probabilistic seismic risk analysis, which identifies regions with a greater likelihood of suffering economic or human losses.

5. ACKNOWLEDGEMENTS

The authors acknowledge the support of Portuguese Science and Technology Foundation (FCT), through grant PTDC/ECI-EGC/7244/2020, project ASSIMI:ATE.

6. REFERENCES

Abrahamson NA, Youngs RR (1992). “A stable algorithm for regression analyses using the random effects model”. *Bulletin of the Seismological Society of America* 82:505–510. https://doi.org/10.1785/BSSA0820010505

Carrilho F, Custódio S, Bezzeghoud M, Oliveira CS, Marreiros C, Vales D, Alves P, Pena A, Madureira G, Escuer M, Silveira G, Corela C, Matias L, Silva M, Veludo I, Dias N, Loureiro A, Borges J, Caldeira B, Wachilala P, Fontiela J (2021). “The Portuguese National Seismic Network—Products and Services”. *Seismological Research Letters* 92 (3): 1541–1570. doi: https://doi.org/10.1785/0220200407.

Crowley H, Despotaki V, Rodrigues D, Silva V, Toma-Danila D, Riga E, Karatzetzou A, Zugic Z, Sousa L, Ozcebe S, Gamba P (2020). “Exposure Model for European Seismic Risk Assessment”. *Earthquake Spectra*. https://doi.org/10.1177/8755293020919429

Edwards B, Fäh D (2013). “A Stochastic Ground‐Motion Model for Switzerland”. *Bulletin of the Seismological Society of America* 103:78‐98. https://doi.org/10.1785/0120110331

Engler, D. T., Worden, C. B., Thompson, E. M., & Jaiswal, K. S. (2022). “Partitioning Ground Motion Uncertainty When Conditioned on Station Data”. *Bulletin of the Seismological Society of America*, 112(2), 1060–1079. DOI: https://doi.org/10.1785/0120210177.

Erdik M, Fahjan Y, Orguz O, Alcik H, Mert A, Gul M (2003). “Istanbul earthquake rapid response and the early warning system”. *Bulletin of Earthquake Engineering,* 1(1):157-163.

Erduran E, Lang D, Lindholm C, Toma-Danila D, Balan S, Ionescu V, Aldea A, Vacareanu R, Neagu C (2012). “Real-time earthquake damage assessment in the Romanian-Bulgarian Border region”. *In: Proceedings of the 15Th World Conference on Engineering,* Lisbon, Portugal.

Fiorini E, Borzi B, Iaccino R (2012). “Real Time damage scenario: case study for the L’Aquila earthquake”. *In: Proceedings of the 15th World Conference on Earthquake Engineering*, Lisbon, Portugal.

Hoshiba M, Iwakiri K, Hayashimoto N, et al. (2011). “Outline of the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0) earthquake early warning and observed seismic intensity". *Earth Planets Space*. 63:547-551.

Jaiswal K, Wald D (2010). "Development of a semi-empirical loss model within the USGS Prompt Assessment of Global Earthquakes for Response (PAGER) System". *In: Proceedings of the 9th US and 10th Canadian Conference on Earthquake Engineering*, Toronto, Canada.

Jaiswal K, Wald D (2013). "Estimating Economic Losses from Earthquakes Using an Empirical Approach". *Earthquake Spectra*, 29:1, 309-324.

Rietbrock A, Strasser F, Edwards B (2013). "A Stochastic Earthquake Ground ‐ Motion Prediction Model for the United Kingdom". *Bulletin of the Seismological Society of America* 103:57 ‐ 77. https://doi.org/10.1785/0120110231

Silva V (2015). "Critical Issues in Earthquake Scenario Loss Modelling". *Journal of Earthquake Engineering*, 20(8):1322-1341. doi.org/10.1080/13632469.2016.1138172

Silva V, Horspool N (2019). "Combining USGS ShakeMaps and the OpenQuake-engine for damage and loss assessment". *Earthquake Engineering and Structural Dynamics*. 48(6):634-652.

Silva V, Marques M, Castro JM, Varum H (2015). "Development and Application of a Real-time Loss Estimation Framework for Portugal". *Bulletin of Earthquake Engineering*, 13:2493–2516. doi.org/10.1007/s10518-015-9743-0.

Wald D, Seligson HA, Rozelle J, Burns J, Marano K, Jaiswal K, Hearne M, Bausch D (2020). "A domestic earthquake impact alert protocol based on the combined USGS PAGER and FEMA Hazus loss estimation systems". Earthquake Spectra 36, 164–182, doi: 10.1177/8755293019878187.

Worden CB, Thompson EM, Baker JW, Bradley BA, Luco N, Wald D (2018). "Spatial and spectral interpolation of ground-motion intensity measure observations". *Bulletin of the Seismological Society of America*. 108, 866–875, doi: 10.1785/0120170201.

Wyss M (2004). "Real time prediction of earthquake casualties". *Disasters and Society - from Hazard Assessment to Risk Reduction*, edited by D. Malzahn and T. Plapp, pp. 165-173, Logos, Berlin.