

Towards time- and state-dependent seismic risk over urban scales

vorgelegt von
M.Sc.
Michael Haas
geb. in München

von der Fakultät VI- Planen Bauen Umwelt
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr.rer.nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Frank Rackwitz
Gutachter: Prof. Dr. Paolo Bazzurro
Gutachter: Prof. Dr. Stefano Parolai
Gutachter: Prof. Dr. Yuri Petryna

Tag der wissenschaftlichen Aussprache: 22. November 2018

Berlin 2019

Abstract

Probabilistic seismic risk assessment involves the appraisal of three components: the seismic hazard due to expected ground motions, the exposure, i.e., assets exposed to the hazard, and their vulnerability with respect to the hazard. Earthquake sequences are a result of a time dependent process, which is complex to consider in a probabilistic framework. It is thus common to reduce sequences to their largest event, the so-called main shock. The occurrence of these main shocks can be modeled assuming a Poisson process with constant average rates. This is convenient from a mathematical point of view, as the connected probability distribution is simple, but of course, a large portion of the seismicity is neglected. As a result of the long inter-event times in such a Poissonian model, it is commonly assumed that potentially damaged buildings are repaired before the next strong event. In reality, the repair of buildings, especially if a large number have been affected, might take several years. As a consequence, buildings are likely to be still damaged when subsequent events occur. For aftershocks, which commonly occur within hours or days after the main event, this is definitely the case. Furthermore, a building's response is not identical for its intact and damaged states. If these effects are not considered within a seismic risk assessment, then the resulting estimates will most likely fail to predict accurate results.

In order to determine if this is the case over urban scales, this thesis sets out to design, develop and test an approach to relax the two assumptions with the aim of developing time- and state-dependent seismic risk models. The accompanying research questions are: 1) if this results in loss estimates that are significantly different from those of a classical model, 2) which components of the model have the strongest influence, and 3) if simplifications can be made to model the complex process.

The approach developed here employs a simulation framework that models full seismic sequences using an epidemic type aftershock model, a Markov-chain damage process, state-dependent fragility models, and various probabilistic repair functions. The approach is tested for the city of Nablus in the West Bank, Palestine. For this purpose, a fully probabilistic seismic hazard model is developed for the region. A multi-source imaging analysis is employed to collect exposure data using remote rapid visual screening. For the identified building types, a simplified analysis is performed to obtain state-dependent fragility models, while the many simplifications applied don't yield reliable absolute losses, they permit the identification of tendencies and answer the formulated research questions.

The results show that a time- and state-dependent model yields significantly higher losses than a classical model, reaching a difference of up to 58% in the examples presented in this work. It is shown that considering the problem only partially is not sufficient. While including fore- and aftershocks, but no time-delayed repair and state-dependency, yields a potentially strong overestimate of losses at low exceedance probabilities, disregarding the state-dependency underestimates the losses over the

whole range of probabilities. Furthermore, considering time- and state-dependency in Poissonian models does not influence the losses significantly, at least for the demonstrated case of Nablus.

It can be concluded that despite the greater efforts required to develop time- and state-dependent models over urban scales, the effects are strong enough to justify the development cost if accurate risk models are intended. However, for observation periods of a single year, it was found that a constant daily repair probability may be sufficient to approximate more complex time-variant repair models. Many additional questions have emerged from the findings of this thesis. These concern the issue of more detailed analyses using more sophisticated components, mainly for the hazard and fragility models. In addition, the framework developed within this thesis can form the basis for a multitude of new research directions in the field of cascading effects. These may include the consideration of multiple hazards, dynamically evolving exposure models, and interactions between them.

Zusammenfassung

Eine probabilistische Erdbebenrisikoanalyse beinhaltet die Bewertung dreier Komponenten: der aus Bodenbewegungen resultierenden seismischen Gefährdung, dem Exposure, sprich Werte, die der Gefährdung ausgesetzt sind und der Vulnerabilität dieser im Hinblick auf die Gefährdung. Erdbeben ereignen sich als Folgen und sind das Ergebnis eines zeitabhängigen Prozesses, dessen wahrscheinlichkeitstheoretische Berücksichtigung kompliziert ist. Deshalb ist es üblich die Erdbebensequenz auf deren größtes Erdbeben, sogenannte Hauptbeben zu reduzieren. Das Auftreten dieser Hauptbeben kann als Realisierungen eines Poissonprozesses mit konstanten Raten simuliert werden. Dies ist aus mathematischer Sicht günstig, da die damit verknüpfte Wahrscheinlichkeitsverteilung einfach zu handhaben ist, allerdings wird durch die Reduktion ein großer Teil der Seismizität vernachlässigt. Da die Zeit zwischen zwei dieser Hauptbeben eines solchen Poissonmodelles verhältnismäßig lang ist, wird für gewöhnlich angenommen, dass die möglicherweise in einem Beben beschädigten Gebäude bis zum nächsten Beben wieder repariert sind. In Wirklichkeit kann die Reparatur von Gebäuden unter Umständen mehrere Jahre dauern, insbesondere wenn eine große Anzahl beschädigt wurde. Daraus folgt, dass im Falle weiterer Beben einige Gebäude wahrscheinlich noch beschädigt sind. Für Nachbeben, die auch innerhalb von Stunden oder Tagen nach dem Beben auftreten können, ist dies mit hoher Wahrscheinlichkeit der Fall. Des Weiteren ist die Antwort eines Gebäudes auf eine Bodenbewegung im beschädigten Zustand nicht identisch zu der eines intakten Gebäudes. Falls diese Effekte im Rahmen einer Erdbebenrisikoanalyse nicht berücksichtigt werden, ist es möglich, dass die resultierenden Abschätzungen daran scheitern akkurate Vorhersagen treffen zu wollen.

Um herauszufinden ob dies über urbane Skalen der Fall ist, wird im Rahmen dieser Arbeit ein Ansatz entwickelt, implementiert und getestet, der beide klassischen Annahmen in Frage stellt, und hin zur Entwicklung von zeit- und zustandsabhängigen Erdbebenrisikomodellen führt. Die damit einhergehenden wissenschaftlichen Fragestellungen sind: 1) ob solch ein Modell Verlustabschätzungen zur Folge hat, die signifikant verschieden von denen eines Modells unter klassischen Annahmen ist, 2) welche Komponenten des Modells den stärksten Einfluss darauf haben und 3) ob zu einem gewissen Grad Vereinfachungen für das komplexe Modell möglich sind.

Der hier entwickelte Ansatz baut auf eine Simulationsumgebung die vollständige Erdbebensequenzen mit Hilfe eines epidemischen Nachbebenmodelles simuliert, einen Markov-Ketten Schadensprozess nutzt und zustandsabhängige Fragilitätsmodelle sowie verschiedene Wahrscheinlichkeitsfunktionen für die Reparatur berücksichtigt. Der Ansatz wird anhand der Stadt Nablus im Westjordanland (Palästina) getestet. Zu diesem Zweck wird ein vollständig probabilistisches Erdbebengefährdungsmodell entwickelt. Zudem wird eine Satellitenbild- und Mobile-Mapping gestützte Bildanalyse multipler Quellen durchgeführt, um mit Hilfe von rascher visueller Gebäudefernerkundung Exposureinformationen zu Gebäuden zu sammeln. Für die so bestimmten Gebäudetypen wird eine einfach gehaltene Analyse durchgeführt,

um zustandsabhängige Fragilitätsmodelle zu erhalten. Obwohl die zahlreichen Vereinfachungen dabei keine verlässlichen Modelle liefern, erlauben diese es Trends zu identifizieren, um die genannten wissenschaftlichen Fragestellungen zu beantworten.

Die Ergebnisse zeigen, dass ein zeit- und zustandsabhängiges Modell signifikant höhere Verlustabschätzungen als ein klassischer Ansatz liefert, der Unterschied erreicht dabei bis zu 58% in den in dieser Arbeit präsentierten Beispielen. Des Weiteren kann gezeigt werden, dass eine nur teilweise Berücksichtigung der Problematik nicht ausreicht. Während ein Ansatz, der nur Vor- und Nachbeben, aber keine zeitverzögerte Reparatur berücksichtigt, Verluste am unteren Ende der Wahrscheinlichkeiten möglicherweise stark überschätzt, führt eine Vernachlässigung der Zustandsabhängigkeit der Gebäudefragilität dazu, dass die Verluste für alle Wahrscheinlichkeiten unterschätzt werden. Außerdem konnte festgestellt werden, dass die Berücksichtigung von zeit- und zustandsabhängigen Komponenten in Poissonmodellen, zumindest für den Fall von Nablus, die Verlusteinschätzungen nicht signifikant beeinflusst.

Es lässt sich schlussfolgern, dass trotz des größeren Aufwandes, der nötig ist um zeit- und zustandsabhängige Modelle über urbane Skalen zu entwickeln, die beobachteten Effekte stark genug sind, um die Entwicklungskosten zu rechtfertigen, wenn akkurate Risikoeinschätzungen gewünscht sind. Für kurze Beobachtungszeiträume von einem Jahr konnte jedoch gezeigt werden, dass eine tägliche konstante Reparaturwahrscheinlichkeit möglicherweise ausreichend ist, die komplexeren zeitvariablen Reparaturmodelle anzunähern. Die Arbeit hat eine große Zahl neuer Fragen aufgeworfen. Diese betreffen etwa die Bestätigung und feinere Ausarbeitung der Trends durch detailliertere Analysen zu den einzelnen Komponenten, das heißt es können anspruchsvollere Modelle entwickelt werden, vor allem für die Erdbebengefährdung und die Gebäudefragilität. Des Weiteren kann die für die Simulationen entworfene Programmrahmenstruktur als Ausgangsbasis für eine Vielzahl an neuen Forschungsarbeiten im Bereich kaskadierender Ereignisse genutzt werden. Diese können etwa die Berücksichtigung verschiedener paralleler Gefährdungen, sich dynamisch entwickelnde Expositionsmodelle und Interaktionen zwischen diesen mit einschließen.

Acknowledgments

This work would not have been completed without the help of a number of people. First of all, I want to thank my supervisor and mentor, Stefano Parolai, who always supported me in every aspect and at any time of the day or wherever he was. Thanks a lot also to my supervisor, Massimiliano Pittore, who supported every step of this work and always provided critical feedback and suggestions. My deepest gratitude to Konstantinos Megalooikonomou who introduced me to the world of structural engineering and never stopping answering my many questions and concerns. Thanks to Sebastian Hainzl for his help understanding and to implement the epidemic type aftershock simulations and his benevolent reminding that I should not work too much. Thanks to Paolo Bazzurro for supporting me with his enormous experience and knowledge during the last miles of the work. Thanks to Yuri Petryna for his critical and constructive feedback regarding my work. I would also like to give my special thanks to Jalal Al-Dabbeek, director of the earth sciences and seismic engineering center of the An-Najah University, who strongly supported us during the field campaigns and whose engineers performed the visual screening that formed the fundamental data set for this study. Without Jalal, this study would not have been possible thank you for your help and friendship. A special thanks also to Kevin Fleming who gave many useful suggestions and fought vigilantly trying to improve my command of English.

Many thanks to Dorina Kroll who helped organize my daily life at GFZ, arranging travels and keeping the monsters of bureaucracy at bay, and thanks also to Susanne Köster who helped out when Dorina was not near. Thanks to my colleagues and former colleagues, Ana-Luisa Ramirez, Dino Bindi, Shahid Ullah, Marc Wieland, Marco Pilz, Bojana Petrovic and Tobias Boxberger for the fruitful discussions regarding a multitude of topics and their friendship.

Thanks to Michael Weber who introduced me to the region and who supported me in any aspect during the DESERVE project. At this point, also thanks to Manuela Nied who coordinated the project.

The ETAS code used in the study was developed and kindly provided by Sebastian Hainzl. The initial implementation of the RRVS tool was developed by Marc Wieland. OpenSees was used to perform the incremental dynamic analysis. The PEER structural performance database was used. OpenQuake was used to generate correlated ground motion fields. Engineers from the An-Najah University (Jalal, Wala, Karim, Anas) analyzed the data used for Nablus. Amotz Agnon from the Hebrew University in Jerusalem, with his outstanding expertise on the subject, kindly guided the review of the historical seismicity for the region.

I want to thank also those people whose efforts made it possible to collect data that was not explicitly mentioned in this thesis, but without them, it would not have been possible to conduct this study and with whom I hope to continue a fruitful

collaboration and friendship in the future. All have helped collecting thousands of images in the Dead Sea area and made it feasible to collect building information for Israel, Jordan and Palestine a seismically very endangered region. In Israel, Tsafir Levi and Jakob Rafael (both of the Geological Survey of Israel) greatly supported the collection of data for Eilat and Tiberias, and many thanks to Tsafir Levi also for his very good introduction into the geology along the Dead Sea fault zone during the field campaign. In Jordan, Ali Sawarieh, Mahmoud Al-Quaryouti (both of the Ministry of Energy and Mineral Resources) greatly supported the preparation and conduction of the field campaign. In addition, Nezar Daoud (Municipality of Madaba) and Abdullah Sarayah (Municipality of Kerak) greatly supported the field campaigns. In Palestine, besides the persons mentioned already, I have to thank Ayman Mohsen, Hatem Alwash (both from the An-Najah national University) for their support preparing for and during my stay in Palestine. Thanks to Husam Jada and especially Safaa Aldwaik from the GIS and IT Department of Ramallah for their support during the field campaign.

Thanks to Abdullah Sarayah who analyzed about 600 buildings in Kerak, providing an insight in the local building stock, thanks a lot for these efforts! Thanks also to Thore Røben who analyzed about 1100 buildings in Israel and 500 in Madaba.

I am grateful for the financial support provided by the Helmholtz Association through the DEad SEar RESearch VEnue, which funded the undertaken work.

Finally, I want to thank my family. I thank my parents for supporting me during all my life, letting me take my own decisions, and find my own ways. Special thanks to my wife Laura who always helped when I needed it, for her understanding for my occasional grumpiness or when I was working late, and in general for keeping her faith in me.

Für Laura und Joshua

Contents

List of Tables	xiii
List of Figures	xv
1. Introduction	1
2. Towards time and state dependent risk	9
2.1. State of the art	9
2.2. Proposed framework	12
3. Poissonian seismic risk on urban scale	15
3.1. Poissonian seismic hazard	15
3.1.1. Seismogenic sources	15
3.1.2. Source-to-site distance	18
3.1.3. Magnitude distribution and rate of occurrence	18
3.1.4. Ground motion and site effects	22
3.2. Fragility and vulnerability of buildings	23
3.3. Exposure	25
3.3.1. Exposure data sampling	26
3.3.2. Building types	27
4. Considering time and state dependency	30
4.1. Modeling full seismic sequences	30
4.2. State dependent damage and time dependent repair	33
4.3. Implementation of the simulation framework	41
5. Application to the city of Nablus	44
5.1. Seismic hazard for Nablus	45
5.1.1. Seismicity model	45
5.1.2. Ground motion models	51
5.1.3. Site-effects for Nablus	51
5.1.4. Resulting hazard estimates	52
5.2. Exposure of Nablus	54
5.2.1. Data collection	55
5.2.2. Exposure model for Nablus	59
5.3. Fragility models for Nablus	63
5.4. Building value and loss ratios	66
6. Simulation results	67
6.1. Resulting loss estimates	67

6.2. Influence of model components on results	71
6.3. Statistical tests	79
6.4. Single year observation period	84
7. Validation of results	92
7.1. Validity checks for the seismicity	92
7.2. Hazard	94
7.3. Exposure model validation	95
7.4. Bias in fragility models	95
7.5. Fully empirical Poissonian risk model for Nablus	97
8. Relevance of the findings and limitations	102
9. Conclusion and Outlook	105
Appendices	111
A. Hazard related	112
A.1. Seismic catalog	112
A.2. Optimized selection of ground motion models	115
B. Exposure related	118
B.1. Simplified satellite analysis	118
B.2. Homogeneous strata for Nablus	120
B.3. Ground based data sampling	123
B.4. Taxonomy and rapid visual screening	123
B.5. RRVS web platform	125
B.6. GEM Taxonomy Codes	127
B.7. Clustering exposure data to reduce complexity	131
C. Fragility related	136
C.1. Building models for Nablus	136
C.2. Eurocode 8 building capacity for moment frame buildings	137
C.3. Building response, deterioration and fragility	146
C.4. Fragility for buildings in Nablus	150
C.5. Comparing fragility method to other approaches	153
D. Simulation input examples	158
D.1. Configuration file	158
D.2. Seismic catalog	160
D.3. Declustered catalog	161
D.4. Depth model	161
D.5. Source mechanisms	161
D.6. Assets	162
D.7. Fragilities	162
E. Computational costs	163
F. Statistical tests	165

G. Glossary	174
H. Software acknowledgment	183
Bibliography	184

List of Tables

5.1.	Year of completeness estimates for 0.5 Mw bins using Burkhard and Grünthal (2009)	45
5.2.	Depth distribution determined from the 100 events closest since 1900 to a set of grid points (longitude, latitude).	48
5.3.	ETAS parameters. Values are taken from Moradpour et al. (2014) except if marked with *, which are derived in the scope of this study.	48
5.4.	Building parameter ranges from different sources and the employed values.	60
5.5.	Building model capacity, price (\pm standard deviation) and share in numbers and value. The model-id is composed by the types typical number of floors and the bounds for the year of their construction.	60
5.6.	Building type with Eurocode EC-8 soil type derived from v_{s30} , and their proportion p [%] in the <i>full</i> exposure model and the <i>sample</i> used for risk estimates herein (red points in Figure 5.14). Note the missing values are not defined in the respective model.	62
5.7.	Lognormal fragility curves for building models on soil type B and C (\pm standard deviation). Only simulated fragilities are shown while the others are calculated using equation 4.20.	65
6.1.	Simulations with 50 year observation periods run with the developed framework under different modelling conditions.	72
6.2.	Average number of transitions in 50'000 years for the different simulations (see Table 6.1); colors rank the number of transitions per damage grade from highest (dark red) to lowest (dark blue).	75
6.3.	Pairwise model combinations whose differences in the loss distribution at different levels of exceedance probability (PoE[%]) are non-significant according to a two-sided Kolmogorov-Smirnov test (KS , p_{KS}) with a significance level $\alpha = 0.05$; d is the size of effect (Cohen 1988) for the mean values.	81
6.4.	Simulation runs for a 1 year observation period under different modelling conditions employing the developed frame work.	86
6.5.	Pairwise model combinations with 1 year observation period whose differences in the loss distribution at different levels of exceedance probability (PoE[%]) are non-significant according to a two-sided Kolmogorov-Smirnov test (KS , p_{KS}) with significance level $\alpha = 0.05$; d is the size of effect (Cohen 1988) for the mean values.	90

7.1.	Average exceedance probabilities of HAZUS damage states ds_1 , ds_2 , ds_3 , and ds_4 for building types aggregated by number of stories. For comparison, the exceedance probabilities as obtained by Grigoratos et al. (2016) for buildings in Nablus using EMS-98 damage states (DS1-5) are provided in brackets.	97
7.2.	Considered Intensity Prediction Equations (IPEs), their performance for three measures, similarity cluster and weights	98
7.3.	Average number of observed damage states within 50 years for the empirical model built in OpenQuake and a model under classical assumptions from the framework ($FR(1)$ in Table 6.1)	101
A.1.	Composition and Sources of the Declustered Catalog.	112
A.2.	Time and distance thresholds, that if not exceeded identify duplicate events in the catalog.	113
A.3.	Magnitude conversion relations applied to the original magnitude types in the catalog.	114
B.1.	Simple contingency matrix	119
B.2.	Short codes used for attributes within this study (please refer to Brzev et al. 2013, for a complete list of GEM Building Taxonomy codes, rows with * are not part of the taxonomy).	127
B.3.	Most similar attribute levels (similarity in braces) for several attributes using the GEM Glossary articles.	135
C.1.	Eurocode-8 soil types and corresponding v_{s30} values (EC-8 Part 1 Table 3.1)	148
F.1.	Statistical test for pairwise model combinations with higher and lower median values at different levels of exceedance probability (PoE[%]). Lavene (Brown-Forsythe; statistic: L and p-value: p_L) test for variance inhomogeneity, Mann-Whitney-U-test for significant shift in distributions (U , p_U), in case of variance heterogeneity t-test on Box Cox transformed samples (T , p_T), Kolmogorov-Smirnov test (KS , p_{KS}) and size of effect (d) for the mean values.	167
F.2.	Statistical test for pairwise model combinations with higher and lower median values at different levels of exceedance probability (PoE[%]). Lavene (Brown-Forsythe; statistic: L and p-value: p_L) test for variance inhomogeneity, Mann-Whitney-U-test for significant shift in distributions (U , p_U), in case of variance heterogeneity t-test on Box Cox transformed samples (T , p_T), Kolmogorov-Smirnov test (KS , p_{KS}) and size of effect (d) for the mean values.	172
G.1.	Glossary of terms used that are not defined within the body of the main text.	174

List of Figures

1.1.	Global seismicity ($M_w 6+$ 1990-2016).	2
1.2.	Risk triangle used to symbolize the strong relationships between hazard, exposure and vulnerability (modified after Crichton 1999).	3
1.3.	Complex damage and repair process during a seismic sequence (top) compared to the simplified assumptions in a classical Poissonian risk assessment (bottom). Note: p is the probability of repair depending on the time since the last damage Δt	6
1.4.	Matching of cumulative losses (blue curves) and cumulative annual premiums (red curve, based on average annual loss). The question of how much larger is the actual loss considering time and state dependency (dashed blue line) with respect to a classical estimate (solid blue line)?	7
3.1.	Different approaches to modeling a seismic source (grey points and planes) as 3.1a area source zone, 3.1b point or 3.1c + 3.1d complex fault geometry. Each of the different source models allows different source-to-site distances d to be calculated, which may consider the distance from the site (blue) to the actual rupture geometry (red planes), or just a point in space (red points).	17
3.2.	Example for observed seismicity and the corresponding maximum likelihood Gutenberg-Richter relation estimated in Haas et al. (2016b, $a = 6.92$, $b = 1.25$).	19
3.3.	Gridded Gutenberg-Richter 10^a values obtained using Frankel (1995) on the presented catalog (gray points) and employing a 0.1 by 0.1 degrees grid, a smoothing kernel with 20 km, and a minimum magnitude $M_w = 4$	21
3.4.	Fragility curves describing the probability of exceeding a building's response corresponding to different damage states (here HAZUS).	24
3.5.	Multi-tier imaging analysis for gathering exposure data, with analysis scale, example images, and obtainable information (modified after Wieland 2013).	26
3.6.	Complex exposure data set collected with the GEM Building Taxonomy, where red squares are unique building configurations and blue circles correspond to all different encountered attributes which are connected to the building configuration through gray lines. Considering that the underlying dataset consist of only 800 buildings, the complexity, i.e., the unique set of building configurations is very large.	29

4.1.	Observed reconstruction ratio with time in years for several globally observed large events and three time dependent lognormal models tested within this thesis: optimistic ($\ln(\mu) = -0.1$, $\ln(\sigma) = 0.4$ [yrs]), conservative ($\ln(\mu) = 1$, $\ln(\sigma) = 0.5$ [yrs]), and pessimistic ($\ln(\mu) = 1.3$, $\ln(\sigma) = 0.5$ [yrs]).	34
4.2.	Observed reconstruction ratios for Kobe (Japan) and its quarters affected by the 1995 earthquake (data from Hirayama 2000); gray circles are the average rate for the whole town.	35
4.3.	Non-linear static pushover curve for a building, and linear unloading after experiencing a displacement of ds_i with unloading stiffness $K_u < K_1$, and residual static displacement D_{rs} (not to scale; modified from Maffei et al. 2004, p. 16).	36
4.4.	Non-linear static pushover curve for a building, and linear reloading from the residual dynamic displacement $D_{rd} < D_{rs}$ with stiffness $K_i < K_1$ and hardening stiffness $K_{hi} < K_h$, after previously experiencing a displacement of ds_i (not to scale; modified from Maffei et al. 2004, p. 17).	37
4.5.	Schematic plot of IDA (Vamvatsikos and Cornell 2002) results (16, 50 and 84% quantiles in green, red, and pink) for a non linear static pushover curve (blue) as returned by SPO2IDA (Vamvatsikos and Cornell 2006) and an indicated lognormal fit (black) describing the fragility for a single damage state ds_i (not to scale; modified from Maffei et al. 2004, p. 22).	38
4.6.	Linear approximation of response yielding a strong overestimation of the capacity and thus an underestimation of the fragility. The resulting fragilities are not realistic, but yield state dependent models suitable for demonstration purposes.	40
4.7.	Monte-Carlo simulation framework to model time and state dependent seismic risk over urban scales. The simulation parameters are defined in a configuration file. The required input are a seismic catalog, a spatial model of depths and source mechanisms, location and type of assets, and associated state-dependent fragility curves (see Appendix D for examples). The simulation starts with generating seismicity in the ETAS module and continues anti-clockwise along the work flow until all simulations as defined in the configuration file are done. Results are independent simulations of the damage and repair history for each building.	42
5.1.	Location of Nablus in the Dead Sea region (triangle), and events ($M_w \geq 4$) contained in the declustered catalog window and harmonized to M_w (circles, stars indicate strong events annotated with their year of occurrence). The main features of the region's tectonics (black solid lines, modified from Salamon et al., 1996) are also shown: RSSZ, Red Sea spreading zone; GS, Gulf of Suez; GA, Gulf of Aqaba; WA, Wadi Araba; DSTF, Dead Sea transform fault; CA, Cyprian arc; EAFZ, East Anatolian fault zone (Figure modified from Haas et al. 2016b)	46

5.2. Spatial probability model for the occurrence of Poissonian seismicity obtained after normalizing the gridded a-values resulting from the Frankel (1995) approach.	47
5.3. Events found in the Global Centroid Moment Tensor database (CMT; Dziewonski et al. 1981; Ekström et al. 2012) for the region of interest and a simplified description of the tectonics (Garfunkel et al. 1981) .	49
5.4. Example for a typical simulated catalog over 50 years; (a) shows the background events obtained from the Poissonian model and (b) the events triggered by the events in (a) according to the inhomogeneous Poisson process as described by the ETAS model. Note that triggered events may include larger events (main shocks) and only events with $M_w > m_{min}$ are plotted here.	50
5.5. Built-up area of Nablus (gray) and v_{s30} values inferred from topography with a resolution of 30 arc seconds ($\approx 800m$; Wald and Allen 2007; Allen and Wald 2009), used as a proxy for the site-effect.	52
5.6. Median hazard curves (solid lines) obtained for a single site in central Nablus (35.268 E, 32.220 N, $v_{s30} = 760m/s$) determined using a 10% bootstrap sampling approach alongside the 5 and 95% quantiles (dotted) and 25 and 75% quantiles (dashed). The left side shows curves resulting from using the ETAS model and the right side using only a Poissonian model.	53
5.7. Median hazard curve (solid line) obtained for a single site in eastern Nablus (35.293 E, 32.206 N, $v_{s30} = 327m/s$) determined using a 10% bootstrap sampling approach alongside the 5, 25, 75 and 95% quantiles (dotted, dashed, dashed, dotted, respectively).	53
5.8. Resulting built-up strata using a pixel-based satellite imaging analysis (Appendix B) distinguishing residential, commercial, and mixed landuse in 2015 and approximate period of urban development ($< 1972, 1972-1992, 1992-2015$).	56
5.9. Route resulting from randomized stratified sampling approach, where randomly sampled nodes from the street network fall into the different strata.	56
5.10. Buildings screened by local engineers via the Remote Rapid Visual screening (RRVS) system using panoramic imagery collected along the path shown in Figure 5.9.	57
5.11. Comparison for exterior wall materials as indicated by the census (on the district level) and the data derived from imagery collected with the GFZ-MoMa and interpreted with the RRVS (city level).	57
5.12. Structural material types in Nablus as observed through the RRVS; Concrete reinforced (CR), unknown reinforcement (C99), unreinforced (CU). Analogous: Masonry (MR, MU, M99); Earthen unreinforced (EU); Steel (S) and unknown (MAT99).	58
5.13. Cumulative density for the year of construction of buildings as indicated in the census (for the whole of Palestine) and as interpreted during the RRVS (Nablus). Even if considering the difficulties to assign building age during visual screening, the differences are quite large.	58

5.14.	Spatially distributed exposure model of RC frame structures in Nablus (gray points). For the risk analysis, only a random sample of 1'000 buildings (red) was used (see Figure 5.9).	61
5.15.	Fragility curves for a 7 storey RC frame built after 2000 with a one-way ribbed slab on EC-8 soil type B, for 5.15a-5.15c different starting damage states (ds_i). Note that grey lines in 5.15a are literature models as obtained from the OpenQuake database (available at https://platform.openquake.org ; accessed on May 2018). 5.15a shows curves for ds_0 and ds_1 as ds_1 is still in the elastic range they are identical. 5.15d shows exceedance probabilities for drifts corresponding to ds_4 from the different starting damage states.	64
6.1.	50 year median loss exceedance probability (PoE) curve for 1'000 simulated buildings in Nablus, determined using a 10% bootstrap sampling (1'000 samples) on 10'000 simulations alongside with the 5, 25, 75 and 95% quantiles (dotted, dashed, dashed, and dotted, respectively). Left side shows curves resulting from using a model with full seismicity (ETAS), delayed repair (DR; $\ln(\mu) = 1, \ln(\sigma) = 0.5$ [yrs]), and state dependent fragility (SD). The right side shows a model using a classical assumption (Poissonian seismicity) and fixed probability ($FR(1)$), i.e., immediate repair.	68
6.2.	Residual between the median loss EP-curves of Figure 6.1 (calculated as median EP-curve Figure 6.1a minus median EP-curve Figure 6.1b). Note the logarithmic scale for the probability.	69
6.3.	50 year median loss exceedance probability (PoE) curves from the time- and state-dependent model (Figure 6.1a aggregated for two individual building types, using a 10% bootstrap sampling approach and 5, 25, 75, and 95% percentiles (dotted, dashed, dashed, and dotted, respectively). 6.3a shows the building type that contributed the most, while 6.3b shows one of the smallest (non-negligible, i.e., > 1 Mio at 1% PoE) contributions.	70
6.4.	Ratio of median loss EP-curves for $ETAS+DR(1,0.5)+SD$ (Figure 6.1a) and value of building stock, aggregated per building type in the subset of 1000 buildings in Nablus, again determined using 10% bootstraps. Left buildings on soil-type B (dark colors) and right on soil-type C (bright colors). Note that not all building types are encountered on soil type C (bright color tones) within this sample.	71
6.5.	Median loss exceedance probability curves, relative to the classical risk model (Figure 6.1b; $FR(1)$ in Table 6.1; relative loss equal to 1). Shown are curves for the first six models listed in Table 6.1 with varying modeling conditions such as modeled fore- and aftershocks ($ETAS$), time-delayed repair with log normal model ($DR(\ln(\mu), \ln(\sigma))$) with mean and standard deviation μ and σ , or repair with fixed probability p ($FR(p)$), and state-dependent fragility (SD).	73
6.6.	EP curve for 1'000 simulated buildings in Nablus, without bootstrapping, calculated from 10'000 simulations for 50 years of seismicity. The gray marked area indicates the part of the curve that is supported by less than 500 simulated years.	76

-
- 6.7. Median loss EP-curves for the Poissonian models, one with only time delayed repair ($DR(1,0.5)$), and one with additional state dependent fragility $DR(1,0.5)+SD$, both relative to a classical risk model $FR(1)$. 77
- 6.8. Median loss exceedance probability curves for different time dependent repair models, relative to losses obtained from a model with the conservative repair model $ETAS+DR(1,0.5)+SD$. The solid curve shows losses for an optimistic model $ETAS+DR(-0.1,0.4)+SD$, the dashed line those of a pessimistic repair model $ETAS+DR(-0.1,0.4)+SD$ 78
- 6.9. Boundary case identified for the difference in loss distributions for 10% PoE between the $ETAS+DR(1,0.5)+SD$ and $ETAS+DR(-0.1,0.4)+SD$ model. With an $\alpha = 0.05$, this difference is considered non-significant using a two sided Kolmogorov-Smirnov test ($p_{KS} = 0.05$). Cohen's d describes the size of effect, i.e., the difference of the sample means normalized by the standard deviation of the pooled samples. 81
- 6.10. Quantile vs quantile plot of the two loss distributions at 10% PoE for the $ETAS+DR(1,0.5)+SD$ and $ETAS+DR(-0.1,0.4)+SD$ model (Figure 6.9), in which case the difference between both is close to significance with $p_{KS} = 5\%$ ($\alpha=5\%$). Red line corresponds to identical quantiles. 82
- 6.11. Loss distributions for the two model pairs which yielded the highest and lowest significant size of effect among all model pairs. The plot on the left shows the highest observed size of effect with $d = 5.88$; on the right with $d = 0.03$ the lowest significant size of effect; p_{KS} : p-value for the Kolmogorov-Smirnov test. 83
- 6.12. Quantile vs quantile plots for the two model pairs which yielded the highest and lowest significant size of effect (see Figure 6.11). 84
- 6.13. Annual loss exceedance probability (PoE) curves for 1'000 simulated buildings in Nablus, without bootstrapping, calculated from 25'000 simulations for a single year of seismicity. On the left curves from a time- and state-dependent model $1y ETAS+DR(1,0.5)+SD$ and on the right side one for classical assumptions $1y ETAS+DR(1,0.5)+FR(1)$. Each curve is based on 25'000 years of simulated seismicity. Hence, the lower probability end corresponds to only 2.5 simulated years and very uncertain. The gray marked areas indicate the parts of the curves that are supported by less then 500 simulated years. 85
- 6.14. Residual between the annual loss exceedance probability (PoE) curves for a $1y ETAS+DR(1,0.5)+SD$ model and a classical $1y FR(1)$ model, both EP-curves calculated without bootstrapping from 25'000 simulations for a single year of seismicity, the lower probability end thus corresponds to only 2.5 simulated years and is very uncertain. The gray marked areas indicate parts of the curves that are supported by less then 500 simulated years. 86
-

6.15. Annual loss exceedance probability curves for models in Table 6.4, considering full seismic sequences <i>ETAS</i> , real mainshocks extracted from the <i>real MS</i> , delayed repair of lognormal shape $DR(\ln(\mu) = 1, \ln(\sigma) = 0.5$ [years], and fixed repair with daily probability p ($FR(p)$), all relative to a classical Poissonian annual risk model ($FR(1)$); relative loss equal to 1). Each curve is based on 25'000 years of simulated seismicity, the lower probability end thus corresponds to only 2.5 years and is very uncertain. The gray marked areas indicate parts of the curves that are supported by less then 500 simulated years.	88
6.16. Annual loss exceedance probability curves as in Figure 6.15, but determined using a 10% bootstrap, hence the probability range is limited at 1%. The gray marked areas indicate parts of the curves that are supported by less then 500 simulated years.	89
6.17. Loss distributions for the two model pairs which yielded the highest and lowest size of effect among all model pairs. The plot on the left shows the highest observed size of effect with $d = 0.54$; on the right with $d = 0.00$ the lowest (non-significant) size of effect; p_{KS} : p-value for the Kolmogorov-Smirnov test.	91
7.1. Mean and standard deviation of earthquake counts over a 50 years period on a regular 0.1 by 0.1 degree grid determined using a 10% bootstrap sampling approach.	93
7.2. Mean and standard deviation of the seismic moment over a 50 years period on a regular 0.1 by 0.1 degree grid determined using a 10% bootstrap sampling approach.	94
7.3. On the left: Poissonian median hazard curve (solid) and 5, 25, 75, and 95 % quantiles for a single site in central Nablus (35.268 E, 32.220 N, $v_{s30} = 760m/s$) as obtained with the simulation framework. The right side shows Poissonian median and 5 and 95% quantile hazard estimates for Nablus from the EMME model (Şeşetyan et al. (2018); available from www.efehr.org) on rock ($v_{s30} = 800m/s$).	95
7.4. Fragility model for an intact 7 floor RC Frame building in Nablus (left) as proposed by Grigoratos et al. (2016) using EMS-98 damage states (DSi), and (right) as determined in this study using HAZUS damage states (dsi; averaged over all 7 floor models).	96
7.5. Hazard curves obtained from the EMS-98 based empirical model converted to PGA using Worden et al. (2012) for a single site in central Nablus (35.260 N, 32.225 E) compared to the mean hazard curve obtained from a Poissonian seismicity model using the simulation framework and averaging over all sites considered within Nablus. Quantiles are indicated as: dotted (5/95), dashed (25/75) and solid (50%)	99
7.6. Percentage of EMS-98 vulnerability classes assigned by surveyors to the buildings screened in Nablus using the RRVS (V99: unknown vulnerability)	100

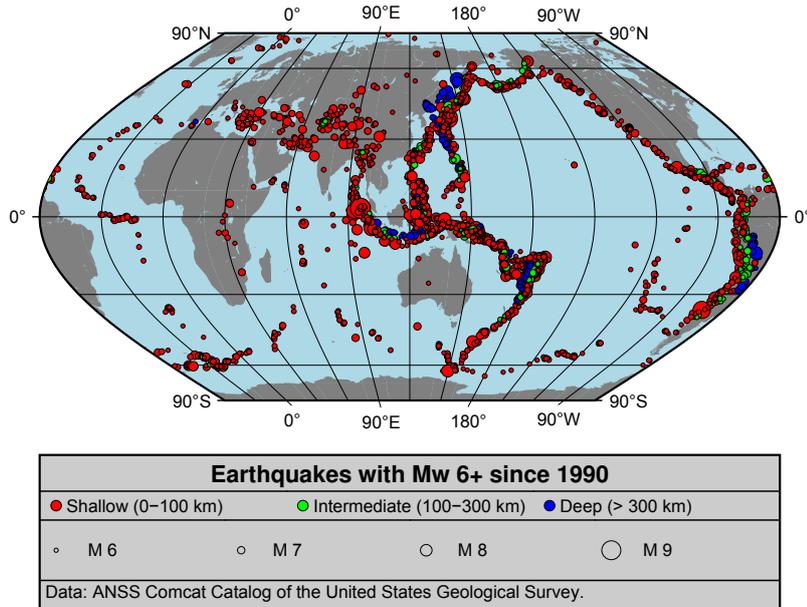
9.1. City at risk: Combining sophisticated socio-economic models and time-dependent hazard simulations will yield a model suitable to explore interactions on the hazard and risk levels. Agent-based models can simulate the dynamics within a city (Waddell 2002), where people are moving, buildings are replaced or new once erected, whereas advanced hazard simulators can be used to model their impact on the city. Finally, a frame work that controls the interactions among all components will allow one to obtain a time- and state-dependent multi risk model.	110
A.1. (a) Sammon’s map (Sammon 1969) of the 10 considered intensity-prediction equations (IPEs); the symbols indicate the cluster that each IPE belongs to, according to a k-means clustering (Steinhaus 1956) of the mean and standard deviations of their estimates over a discrete parameter space. (b)-(f) Estimated mean (curves) and 1σ (shaded) intensity attenuation resulting from each IPE in each cluster for an M_w 7.25 event at 10 km depth against historical distance intensity pairs for events with M_w 7.0–7.5. IPE IDs are (see also Table 7.2) 1, Allen et al. (2012); 2, Bakun (2006); 3, Bindi et al. (2011); 4, Chandler and Lam (2002); 5, Dowrick and Rhoades (2005); 6, Gasperini et al. (2010); 7, Le Goff et al. (2014); 8, Pasolini et al. (2008); 9, Sørensen et al. (2009); 10, Sørensen et al. (2010). Figure modified from Haas et al. (2016b)	116
B.1. Training and testing set derived for Nablus for the five classes residential, commercial, mixed built-up, soil, and vegetation.	120
B.2. Example of training features (red polygons) for (a) residential, (b) commercial, (c) mixed, (d) vegetation, and (e) soil; visually derived from Google satellite imagery (Google, DigitalGlobe, Geoeye 2017).	121
B.3. Resulting classification using REM SatEx for residential, commercial, mixed built-up, soil, and vegetation classes (see Figures B.2).	122
B.4. Screenshot of the RRVS web tool, composed of an overview map (top right) showing a set of buildings (polygons) and available panoramic images nearby (points), a viewer for these images (top left) and the data entry form for attributes based on the GEM Building taxonomy.	126
B.5. Subset of the resulting similarity matrices for the attribute ‘Material Type’. (a) the cosine similarity approach using the short textual descriptions as in Brzev et al. (2013), and (b) attribute articles from the GEM Glossary. For a list of the attribute value codes used in this thesis, see Table B.2	133

C.1.	Cross sections of a column with height H , width b , and depth h_c . Longitudinal reinforcement (gray), and the area (striped) confined by the stirrups (black). Figure C.1a shows a cross section in the $x'y'$ -plane with the compression zone depth x to the neutral axis (<i>N.A.</i>), the thickness of the concrete cover c , and the effective depth d . Figure C.1b shows a column cross section in $y'z'$ -plane with the stirrups (vertical black lines), the longitudinal reinforcement (horizontal black lines), and the area (striped) confined by two stirrups with spacing s . Lap splicing is indicated with the splice length L_s . The red line at an angle α indicates a potential crack due to shear, where $d_2 = d$ in case $\alpha = 45^\circ$	138
C.2.	Loaded cross section, strain diagram and stresses. To iteratively determine the compression zone depth, the concrete is split in layers with height dy (Figure C.2a). Curvature Φ (Figure C.2b) at yielding from the assumption that the tensile reinforcement with maximum distance l_i from the neutral axis reached a strain ϵ_s equal to its yielding strain. From this the tensile stress T_i (Figure C.2c) is calculated for each layer of reinforcement, considering that the parabolic compression stress distribution can be approximated with a rectangular distribution with depth $0.8x$. Calculating the strain using the curvature Φ for each concrete layer, and from these the compression stresses C_i , considering a strength and strain change in confined sections (striped area; see also Figure C.1b). The desired compression zone depth x is found, if the tensile and compressive forces calculated from the C_i s, T_i s and affected areas yield the applied load N	139
C.3.	Ground acceleration recorded during an event at a station with soil-type B (EC-8) in Greece (HI.PAT1..HN2.D.19930714.123148.C.ACC.ASC) and response of an ESDOF system once approximating an intact building and once a extensively damaged version (ds_2 in equation 3.10 of a building. Please note that in the actual simulation not necessarily the same record is chosen for both systems as their eigen period might not match the EC-8 criteria.	149
C.4.	Example for a) maximum drift results for various ground motions (PGA) for a single ESDOF (308 simulations) and b) a fragility curve fit to these drift values for going from damage state ds_2 to the next damage state $ds_3 = 0.021$	150
C.5.	Disaggregation in 0.5 Mw and 10km distance (R) bins of the mean hazard curve on rock for Nablus obtained using the hazard model by Haas et al. (2016b) and a single GMPE (Campbell and Bozorgnia 2014). Only M-R combinations with at least 0.1 % chance of exceeding the mean hazard level with 475 years average return period are shown.	151
C.6.	Recordings selected according to Eurocode-8 requirements from the ESM database (Luzi et al. 2015) employing REXELite Iervolino et al. (2011).	152

C.7. Hysteretic behavior observed during the experiment for an event with spectral acceleration of about 1.1 g at 0.4s by Pinto et al. (2002) (left), capacity curves resulting from the model developed here, in both directions, and the idealized bilinear capacity model (dashed line) fitted via least square residuals to the weaker directions' curve (right).	154
C.8. Fragility curves for the intact ICONS building on EC-8 soil types B and C. Gray curves indicate estimates for RC Moment Frame structures available from the OpenQuake database (https://platform.openquake.org/ ; accessed May 2018).	155
C.9. Increasing probability of exceedance for the ICONS buildings starting from different damage states on EC-8 soil type B.	156
C.10. Fragility curves for the damage states <i>moderate</i> , <i>extensive</i> , and <i>collapse</i> for the ICONS building using SPO2IDA (Vamvatsikos and Cornell 2006)	156
C.11. Performance points, determined at the spectral displacement values corresponding to the HAZUS damage states, using EC-8 elastic and inelastic response spectra, the latter depending on the ductility of the building capacity at the performance point.	157
E.1. Computation time depending on the number of simulations and number of threads (left) and number of buildings (right). While an increase of simulations increases the computation time linearly, increasing building number causes a non-linear increase. The number of threads linearly reduces the computational time with a small overhead, which is only relevant for small numbers of simulations.	164

1. Introduction

Earthquakes have long been considered an inexplicable disastrous event. Often, the occurrence of earthquakes was associated as being fated, or as a punishment by some divine entity (Agnew 2002). In Europe, this understanding lasted until the 18th century, although figures such as Aristotle had proposed a natural cause, which he imagined as being due to winds in underground caverns. Attempts to measure earthquakes are documented as early as year 132 C.E., when in China a device to determine the direction of an earthquake's origin was built by Zhang (Agnew 2002). The first network of seismic instruments was established in Italy by Bertell and de Rossi in 1873, and a few years later the first instrumental record of ground shaking was obtained by Ewing in Japan. Afterwards, Milne, and later Omori, started to routinely report Japanese earthquakes. Earthquakes were monitored only locally until Rebeur-Paschwitz accidentally obtained the first teleseismic record of a Japanese earthquake in 1889 during a gravity experiment in Potsdam (Agnew 2002). Although as global instrumental coverage increased and it became clear that earthquakes tended to occur along belt-like structures (see Figure 1.1), this was not put into a proper perspective until the 1960s (Shearer 2011, p. 5). With his observations of matching African and South American coast lines, Alfred Wegener (1915) supported earlier ideas of a dynamic planet. A decade later, Arthur Holmes (1928) was the first to propose a theory in the direction of modern understanding of mantle convection. This process, the driver of plate tectonics, is the underlying mechanism responsible for most earthquakes. Only decades later was mantle convection accepted by the scientific community, and earthquakes finally came to be understood as a natural phenomenon with scientifically explainable causes. The introduction of the elastic rebound theory by Reid after the 1906 San Francisco earthquake provided the modern understanding of the development process of earthquakes (Shearer 2011, p. 13). According to this theory, stress is accumulated along preexisting fault planes, which at some point is large enough to fracture the rock. Consequently, the two adjacent segments of rock material slip along each other to reach a new stable position. This process releases large amounts of energy, mainly in form of friction and deformations, but also as seismic waves (Shearer 2011, p. 277). Although the involved processes are quite well understood, our knowledge is still not sufficient to predict earthquakes, neither by size, location nor time (Shearer 2011, p. 310). Many reasons for this are routed in the long duration of the processes involved, as plates move with rates of the order of a few centimeters per year, and the slip of a large earthquake, for example a moment magnitude M_w 7 event, is of the order of meters (Wells and Coppersmith 1994). This long process is contrasted by the comparatively short periods for which we have recordings. In addition, our knowledge is limited as the means available to measure solid earth processes are limited. These involve measurements on the surface or at depths that are at the order of meters or few

Figure 1.1.: Global seismicity ($M_w 6+$ 1990-2016).

kilometers.

While the study of earthquake processes is relevant to understanding nature and the causes of earthquakes, it is not these processes that damage the built environment. The phenomenon, i.e., the rupture itself only rarely reaches the surface and is usually limited to a relatively small area. The part of the energy released by an earthquake that is dangerous to human society is usually the seismic waves and resulting ground motions. In addition, the rupture and ground shaking can also trigger a multitude of secondary effects, like tsunamis, landslides, or liquefaction, which in many cases can cause even greater damages than the shaking (Super Cat; e.g., RMS 2005).

Since earthquakes can't be predicted, humanity is currently left with only two options to mitigate the ground motion related risk: early warning of incoming seismic waves and seismic risk assessment. While the former is a reactive tool, i.e., includes early detection, rapid signal processing and alerting, the latter is a process that fosters preparedness and prevention. Traditionally, risk assessment is separated into the analysis of three components, the hazard, the exposure and the vulnerability (Crichton 1999, Figure 1.2;). In the case of seismic risk assessment, the first component, i.e., seismic hazard, is the expected level of ground motion at a site. Exposure refers to assets exposed to the hazard, primarily structures, but may also include the occupants, contents, or goods. The last component, the vulnerability, describes the response of the exposed assets to the ground motion in terms of losses. The seismic risk is therefore the convolution of these three components, which are strongly related and influence the risk. The risk triangle in Figure 1.2 shows a balanced view of the influence each component has on risk. However, in reality, the triangle can have any shape, i.e., the contribution of each component to the overall risk may vary greatly. Well prepared societies exposed to a high hazard, for example, can have lower risk than ill prepared, highly vulnerable, societies with lower hazard. Generally, one distinguishes between deterministic and probabilistic



Figure 1.2.: Risk triangle used to symbolize the strong relationships between hazard, exposure and vulnerability (modified after Crichton 1999).

risk analyses. While probabilistic estimates give an integrated picture, deterministic, i.e., individual earthquake scenarios, are more intuitive and can provide more specific information. However, throughout this study, only probabilistic seismic risk assessment is discussed.

For a fully probabilistic seismic risk assessment each component is treated probabilistically. The hazard at a location is described as the probability of a level of ground motion being exceeded within a given period. For the response of an exposed structure to this ground motion, one distinguishes between fragility and vulnerability. Fragility probabilistically describes the physical consequences, i.e., damage, on the structure caused by the motion. It is common to divide the resulting damage range, typically increasing with ground motion intensity, into several levels of damage severity, from small cracks in the fabric, until complete structural failure. Such damage causes direct and indirect losses. Social consequences might include occupants suffering injuries of varying severity, while economic losses may arise as damaged structures require repairs, building contents may need to be replaced, or structures and infrastructure rendered unusable in turn hamper business operations, whereas vulnerability functions directly describe the expected loss at a given level of ground motion. Originally, risk assessment was intended for seismic design of structures (Cornell 1968). However, seismic risk is also of great interest to governments, urban planners, and civil protection authorities, as well as financial risk bearers such as the insurance industry. With this broader view, assessments are usually carried out for a large population of buildings rather than individual buildings. Unfortunately, inventories of the location of structures, as well as their structural and non-structural features, are rarely kept over such large scales. Thus, this data must be collected, with a common procedure being to disaggregate census data, complemented or substituted by the collection of in-situ data samples. As usually no full census of buildings is conducted, this introduces uncertainties in the location and number of structures, as well as in the features that determine their fragility. As a result, the exposure is also a probabilistic component. All of the underlying parameters are subject to varying degrees of uncertainty. In the most general way,

one can distinguish the uncertainties as being either aleatory or epistemic uncertainty. While the first is attributed to the stochastic nature of the processes, the latter describes the uncertainty stemming from our lack of knowledge.

In order to perform a probabilistic seismic hazard assessment (PSHA; Cornell 1968) for a location, several factors need to be determined. These are 1) the potential earthquake sources in the surrounding, 2) the rate with which these produce earthquakes, 3) the magnitude of these, 4) the associated source-to-site distances, and 5) the level of resulting ground motions. Potential sources are commonly identified from historical observations and geological investigations. Once these sources are identified, the frequency-magnitude distribution of the earthquakes generated by these sources is estimated from observed seismicity and strain rates on faults, if available. According to the seismic rebound theory, the earthquake process is cyclic with an average constant rate, as the underlying geological processes are relatively constant. Indeed, observed seismicity shows that large earthquakes occur at a fairly constant rate. However, earthquakes also occur as sequences, meaning larger events are accompanied by many smaller events clustered in time and space. Commonly ascribing the largest event of such a sequence as the main shock, the accompanying events are termed fore and aftershocks. The sequential behavior stems from the fact that an earthquake is not only releasing stress along its rupture plain, but is also changing stress levels in its surroundings. This occurs dynamically, i.e., through the shaking, but also statically through permanent deformations. Such changes can trigger additional earthquakes (Felzer and Brodsky 2006), leading to cascading behavior. While shortly after the main event, many aftershocks follow rapidly, this process stabilizes after some time and slowly fades (Omori 1894). As Utsu and Ogata (1995) have shown for the Nobi (Japan) earthquake in 1891 and recorded events from 1900 till 1991, sequences can span over 100 years. The process shows fractal, i.e., self-similar, behavior and can thus be found on any scale (Kagan and Knopoff 1987). Thus, in order to describe the rate of earthquake occurrence within such a sequence, a time-variant probabilistic model has to be used. As this is rather complex, analysts usually argue that the main damage caused by each individual earthquake sequence can be attributed to the largest event in the sequence. Identifying these events should then yield a set of mutually independent events that may be attributed to tectonic loading and hence show constant average rates over the considered time period. If the seismicity can be reduced to these events, they can be modeled as realizations of a simple homogeneous Poisson process with a constant, e.g., annual, rate μ . The time Δt between two events can then be sampled from the probability density function of a Poisson distribution as:

$$p(\Delta t) = \mu e^{-\mu \Delta t} \quad (1.1)$$

Based on the decision to model only mainshocks, it is usually assumed that the inter-event time is sufficient to repair structures to their intact state. Both assumptions are what will be referred to as classic risk assumptions in the remainder of this thesis.

However, by removing fore and aftershocks, this classical approach is neglecting large portions of the seismicity. This might prove dangerous as ground motions caused by such events, although potentially smaller than the ones caused by the mainshock,

might still be able to cause further damages. In addition, aftershocks might occur closer to populated areas and despite their smaller magnitudes, may cause stronger ground motion at the site than the main shock (APSHA; Wiemer 2000; Yeo and Cornell 2005). If a seismic risk model includes the complete seismic sequences, then the second assumption of sufficient repair times becomes questionable. Even if only main shocks are considered, the time to repair structures might take years, depending on a multitude of factors, thus the inter event time might not be suffice to repair the complete building stock. Even if the latter should be negligible in case of a Poissonian model with relatively low rates, if one considers full seismic sequences, the time between strong events can be only hours or days. A consequence of such time-delayed repair is that buildings are damaged when the next ground motion occurs and the response of a damaged structure is usually more pronounced than its response when intact (Aschheim and Black 1999). Figure 1.3 illustrates the problem for a sketched example sequence, while the actual damage and repair process can be rather complicated the classical assumption reduces the sequence to a single event and assumes immediate repair.

Both of these risk assumptions can yield a bias in the seismic risk estimates, which need to be as accurate as possible. Civil protection authorities may use risk estimates on urban scales to build an understanding of critical regions within a town. Such models allow them to determine regions which will be strongly affected, i.e., where most people in need of help would be located. Sufficiently accurate models also allow one to judge more detailed questions, such as the distribution of debris, which is a very relevant point as, for example, important roads might be closed or facilities such as hospitals hard to reach. An accurate estimate of the seismic risk is thus paramount to guaranteeing an optimal emergency response. However, accurate risk estimates are not only relevant from an information perspective, but also from an economic point of view. Earthquake damages are among the most costly natural disasters able to cause large losses, for some countries, the resulting losses may exceed a multiple of their gross domestic product. Since this would leave countries, companies or individuals regularly unable to cope with such losses, economic earthquake risk is at least partially transferred to the global financial market through the insurance industry. Distributed over several insurers, the large losses caused by an earthquake can be covered. As with any risk insurance, however, the underlying idea is that through regular payments of annual premiums, a client is guaranteed by the insurer that in case an earthquake causes economic losses, at least part of these are covered by the insurance contract. In order to estimate the required annual premiums to cover annual losses insurance and reinsurance industry commonly base their annual premium calculations on seismic risk assessments. Usually these are based on the expected average annual loss. Calculating loss exceedance probabilities and turning them into average return periods allows these industries, to balance expected costs and annual payments.

Figure 1.4 shows a sketch of such a return period-based loss curve compared to a cumulative annual premium curve. Ideally, both curves would align, usually however in order to account for uncertainties and to pay for the risk that the insurance company takes, e.g., that a large event occurs before enough premiums have been collected, and to secure a revenue for the insurer, the premiums are higher than the

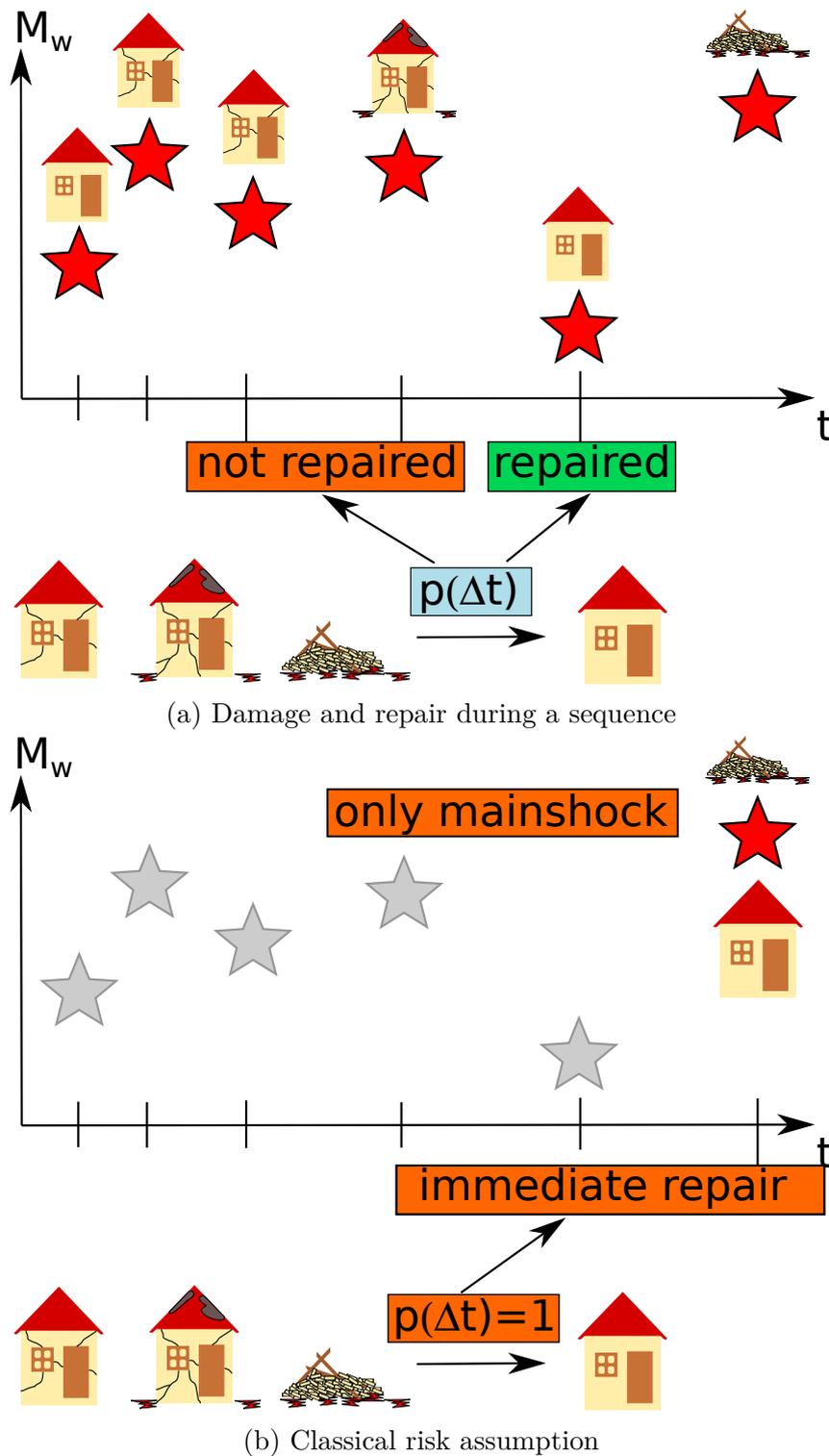


Figure 1.3.: Complex damage and repair process during a seismic sequence (top) compared to the simplified assumptions in a classical Poissonian risk assessment (bottom). Note: p is the probability of repair depending on the time since the last damage Δt .

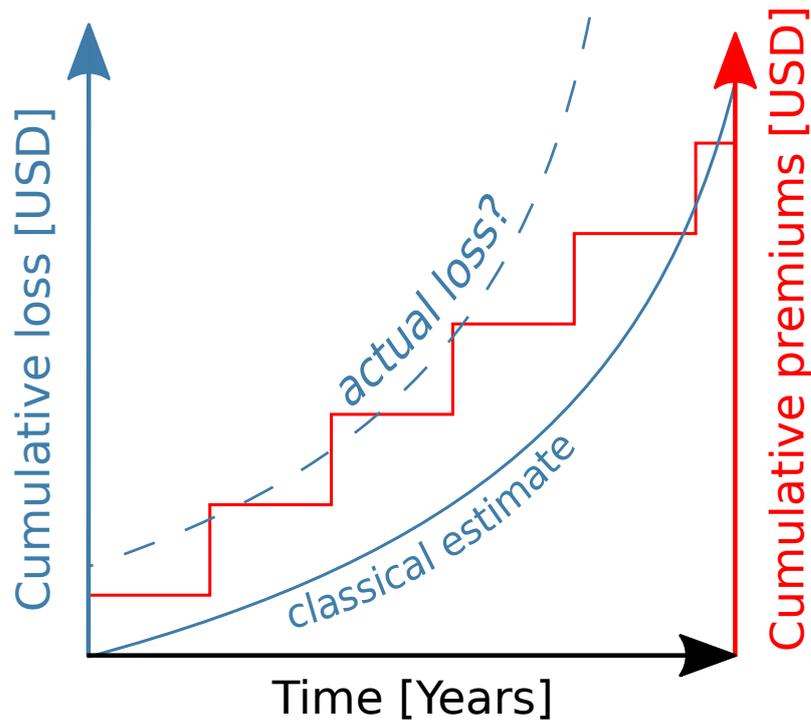


Figure 1.4.: Matching of cumulative losses (blue curves) and cumulative annual premiums (red curve, based on average annual loss). The question of how much larger is the actual loss considering time and state dependency (dashed blue line) with respect to a classical estimate (solid blue line)?

cumulative loss curve.

The purpose of this study is to develop a simple framework that allows the assessment of the relevance of time-dependent hazard, time-delayed repair, and damage state dependent fragility for seismic risk assessment over urban scales. The models will be based on simplified components for the hazard, exposure and vulnerability, intended to test the framework and to determine if more thorough assessments are justified.

Following this brief introduction (**Chapter 1**), **Chapter 2** will introduce how time- and state-dependent risk assessments can be carried out over urban scales. Building on results of previous studies in this direction, an approach is proposed that can be implemented within a probabilistic simulation framework.

For the proposed approach, various components have to be derived. The approach here is to use only simplified components which can be replaced with more sophisticated models in an actual application. **Chapter 3** introduces all components required to develop a seismic risk model at urban scales. Introducing Poissonian seismic hazard assessment, large-scale exposure modeling based on rapid environmental mapping and multi-source imaging, and state-independent fragility and vulnerability models.

Building on these components, **Chapter 4** describes the steps to be taken to develop a non-Poissonian seismic hazard model and how to consider the complex damage

and repair process in a probabilistic manner. Finally, a simulation framework is proposed that is able to incorporate all these components and yield time- and state-dependent risk estimates at urban scales.

This study was undertaken within the framework of the DEadSEa Research VenuE (DESERVE), a Helmholtz funded virtual institute that dealt with environmental challenges within the region around the Dead Sea. **Chapter 5** thus demonstrates the derivation of components for the framework for the city of Nablus in the West Bank, Palestine. Although Nablus is prone to seismic hazard, a seismic building code has been introduced only recently. Many buildings of the relatively homogeneous building stock are thus without seismic design.

Chapter 6 presents the simulation results obtained for Nablus. The results are compared to those using classical risk assumptions which can also be obtained from the simulation framework. Furthermore, a non-exhaustive sensitivity analysis is run comparing risk estimates under various considerations. These include the negligence of individual components in order to determine their influence, and the variation of the repair model. Statistical tests of the results determine significance of the results and their effect on the loss estimates.

As many components are potentially oversimplified and might yield unreliable risk estimates, **Chapter 7** tests and compares the components, and develops a fully empirical model for comparison. The empirical model can not consider the time and state dependency, but allows to judge validity of the framework's simulation results under classical assumptions, as it includes epistemic uncertainty in the hazard and statistically stable fragility models.

Chapter 8 discusses the relevance of the results and under what limitations conclusions can be drawn from the results.

These are summarized in **Chapter 9** and a brief outlook regarding future research, that could build and expand up on the results of this work, is given.

Seismic risk assessment is a highly interdisciplinary process involving seismology, engineering, geography, and statistics. The terminology and concepts are thus many, hence to ease readability some terms are not discussed in detail in the thesis, but a glossary is provided in **Appendix G**.

Parts of this work also have been described in Haas et al. (2016b).

2. Towards time and state dependent risk

As outlined in the introduction, the aim of this thesis is to search for, propose, and test an approach to relax two assumptions very commonly used in seismic risk assessments, namely that sufficiently accurate seismic risk estimates can be obtained by considering only main shocks, and the inter-event time is sufficient to repair structures to their intact state. Targeting a risk formulation without these strong constraints requires one to model seismicity as a time-dependent phenomenon and the response and associated fragility of buildings as being damage state dependent, which in turn depends, at least, on the time it takes to repair a structure. Within this thesis, an approach is developed that allows one to:

- construct time- and state-dependent risk models over urban scales,
- obtain differences in the resulting losses with respect to a classical risk assessment,
- modify and test the influence of individual model components and decisions,
- evaluate the reliability of model simplifications.

In the following, a framework to answer these questions is developed, integrating and expanding various partial solutions that have previously been proposed.

2.1. State of the art

Both of the classical risk assumptions have been targeted, mostly separately, in several past studies, however, no complete framework has yet been proposed that allows time- and state-dependent seismic risk over urban scale. Building on such studies however, a framework can be developed from the individual components. In the following, few important milestones regarding the analysis and modeling of earthquake clustering, the assessment of the related seismic hazard, damage accumulation and time- and state-dependent risk assessment are discussed. Please note that the works presented here are only a selection and by no means exhaustive.

The clustering behavior of earthquakes in time and space was recognized as early as the late 19th century. One of the first set of complete earthquake sequences was

analyzed by Omori (1894) for three earthquakes in Japan. He found that the rate of events following a strong earthquake decreases over time and that the activity depends on the magnitude of the initial shock. For the decay of the rate, he found a simple "nearly satisfactory" (sic) relation that depends only on time t and two constants, K and c (Omori 1894):

$$n(t) = \frac{k}{c + t}.$$

where K is the productivity of a sequence and c the time delay in the on-set of the sequence. This law, commonly known as the Omori law, was modified by Hirano (1924) who added an exponent p to obtain a power-law decay of the aftershock frequency:

$$n(t) = \frac{K}{(c + t)^p}.$$

where p is commonly close to 1. It was also recognized that mainshocks, fore and aftershocks follow the same magnitude frequency relation as main shocks (Gutenberg and Richter 1944).

Cornell (1968) introduced the concept of PSHA using a constant rate Poisson process to model earthquake occurrence. However, he pointed out event then that the Poisson assumption is not consistent with the elastic rebound theory, or able to capture seismic sequences, but that this could be overcome using other, more generic process models. Such aftershock models based on Markov-chain models, referred to as restricted trigger models, were proposed by Aki (1956) and later Vere-Jones (1966), but have not proven adequate for obtaining realistic sequences (Vere-Jones 1966). A complete statistical model that considers the cascading behavior of aftershocks was formulated by Ogata (1988), referred to as the Epidemic Type Aftershock model (ETAS). It relies on an inhomogeneous Poisson process and until now, the Omori law on which it is built remains the best statistical model for aftershock decay (Helmstetter and Sornette 2002).

In order to consider aftershocks in seismic hazard estimates, a number of different approaches have been proposed. Yeo and Cornell (2005) and Yeo and Cornell (2009a) developed an approach for aftershock probabilistic seismic hazard analysis (APSHA) for post-earthquake safety evaluation. The approach is similar to standard PSHA analysis, but the earthquake rate is replaced by a time dependent daily rate based on the modified Omori law. Recently, Iervolino et al. (2014) proposed an approach that combines PSHA and APSHA to sequence-based seismic hazard analysis (SPSHA) using an ETAS model. Both approaches yield, as with a classical PSHA, exceedance probabilities for levels of ground motion over a given period of observation. In contrast to these Beauval et al. (2006), proposed a simulation based method, also based on an ETAS model, but instead of providing a closed form solution, it simulates independent realizations of seismicity within a Monte Carlo sampling framework (Ulam et al. 1947; Metropolis and Ulam 1949) for a given period of interest. At a given site, the hazard levels with different exceedance probabilities are then calculated from the quantiles of the maximum ground motions simulated within a period.

The insurance industry is aware of the fact that risk assessment based on classical seismic hazard estimates might miss damaging events. Consequently, a common way to include these is to consider full sequences as hazard input to the risk models, but to simplify the problem as such that exposed structures are not accumulating damage in subsequent events (Jiménez-Huerta 2009).

In contrast, damage accumulation has been increasingly acknowledged within the earthquake engineering community, as its scope has changed within the last 100 years. Until the second half of the 20th century, earthquake engineering mainly dealt with constructing structures which provide life safety during severe events. For a long time, this was mainly based on experience, but interest in scientific approaches grew with the occurrence of strong events in the late 19th and early 20th centuries, especially in Italy, Japan, and the United States of America (Bozorgnia and Bertero 2004, p. 2). Over the 20th century design aims increased (Bozorgnia and Bertero 2004, p. 10) from only life safety to more complex concepts of damage and harm minimization to performance-based engineering (SEAOC 1995). Additionally, Aschheim and Black (1999) introduced concepts to acknowledge that structures damaged in a previous event can be more susceptible to damage in a subsequent event.

Cumulative damage effects and increased fragility are important factors when determining the safety of mainshock damaged structures (Bazzurro et al. 2004; Luco et al. 2004). Several studies have focused on the probabilistic monitoring of structures after a large seismic event (e.g., Kunitani and Takada 2008; Song et al. 2014; Chioccarelli and Iervolino 2015). Additionally, the complex damage history and behavior of structures is also important from an economic perspective. Therefore, Yeo and Cornell (2005) and Yeo and Cornell (2009a) developed an approach for a building's lifetime loss estimates considering aftershocks (using, e.g., an inhomogeneous Poisson model) in combination with a Markov Chain model for damage accumulation. Yeo and Cornell (2005) suggest the use of time-dependent repair functions using exponentially distributed repair times with relatively short duration, distinguishing between worst and best cases. The best case is a minimum of only 6 days after the main shock to repair a slightly damaged structure, and a maximum of 360 days for severely damaged structures in the worst case. Clearly, the focus lies on a cost optimized decision regarding evacuation or repair of an individual structure during a sequence, based on the probabilistic assessment and is not intended for large scale models. For the life-time loss of a building, Yeo and Cornell (2009a) found an increase of about 25% when compared to a main shock based risk assessment.

Several recent studies have focused on the development of fragility models to account for cumulative damage on a single building scale. For example, Trevlopoulos and Guéguen (2016) propose a damage state-dependent vulnerability assessment based on period elongation as a measurable manifestation of stiffness degradation in damaged concrete buildings of different types. Similarly to Yeo and Cornell (2009b), they used a Markov Chain model and find a probability increase of up to 107% to exceed a period elongation of 60% when considering aftershocks with respect to considering only the mainshock of a sequence. This increase varies considerably between different building types, ranging from 10 to 207%. Regarding the exceedance

probability for the highest damage state Trevelopoulos and Guéguen (2016) found an increase with increasing initial damage state which is below 5% for low rise low code buildings and up to 30% for mid rise low code buildings, but the latter has also the lowest elongation rate increase.

Expanding on the findings of Yeo and Cornell (2009b), Iervolino et al. (2016) proposed a generic framework for stochastic modeling of seismic damage accumulation using a Markov chain that considers the damage state of a structure at the time of an event. The individual transition probabilities from one damage state to another are obtained using a single degree of freedom (SDOF) model with two different response models: a bilinear (elastic perfectly plastic) and an alternative evolutionary pinching model. Both are evaluated in terms of maximum experienced drift values and damage is accumulated through residual displacements. The probabilistic models resulting from dynamic analysis are state dependent and are turned into a unit time transition matrix by multiplying, annual rate of ground motion levels at a site with the individual transition probabilities given a level of ground motion. The authors suggest to choose a unit-time interval that guarantees the probability of having more than one event within the time interval is negligible. This means assuming a constant rate of seismicity. Iervolino et al. (2016) proposed to use a time-variant ground motion rate for the evaluation of an aftershock sequence following a mainshock. To consider repairs, Iervolino et al. (2016) suggest following Cesare et al. (1992) who introduced a fixed probability to repair a structure within the considered time interval in the transition probability matrix. This framework does not allow the consideration of long-term seismic risk as the repairs and damage cannot be easily combined, since the probability of repair depends on the time passed since the last damage occurred, which is itself varying with time.

Recently, Tesfamariam and Goda (2017) presented a study expanding on their assessment in Tesfamariam and Goda (2015) for non-code conforming RC buildings in British Columbia, providing a framework for sequence-based loss assessment for a set of buildings distinguishing different types of earthquakes. Only individual buildings are considered and no repairs are considered. With a median collapse probability increase by 13%, they found the strongest influence of aftershocks were on 2 story low rise buildings and the highest losses for sequences arising from interface events in British Columbia. They found no significant increase for larger structures.

Performing a sequence based risk assessment on regional scale in Umbria Italy, Papadopoulos and Bazzurro (2018) recently showed that the choice of parameters for an ETAS seismicity can affect the estimated losses significantly and must be considered carefully.

2.2. Proposed framework

In order to keep the individual components of risk simple and easily modifiable, this study opts to employ a simple, but consistent simulation framework rather than

expanding existing closed form solutions (e.g., Iervolino et al. 2016) to the needs at hand.

Given the extent of a city, a core component of such a simulation framework is an exposure model defining the locations and features of individual buildings. In case such detailed information is not available, a representative dataset can be collected by employing rapid environmental mapping and multi-source imaging analysis (Wieland et al. 2012). From these data, building typologies, their relative frequency, and their features relevant to seismic response can be determined. Such a model will be stochastic in nature and various strategies can be followed to generate realistic exposure models. This thesis compares the gainable information of such an approach to information that can be obtained from national statistics.

The framework will build upon an ETAS model that generates full seismic sequences within an observation period as suggested by Beauval et al. (2006). The benefit of a simulation-based framework is that the resulting ground motions can not only be used to calculate the probabilities of ground motion occurrence, but they also directly provide realistic ground motion histories for a site. The framework proposed here will show how this can be integrated with other time-dependent components, such as the repair process. Ideally, the ground motions consider local amplification effects and the spatial correlation of ground motion. This can be at least partly fulfilled by employing state of the art seismic hazard software for the ground motion estimation (e.g., Pagani et al. 2014).

If the ground motion is strong enough to damage structures, the time required to repair individual buildings must be considered. Several alternative models will be developed and proposed here. Meanwhile, the damaged structure is potentially subject to further ground motion. The response of damaged buildings is complex to model and requires well-constrained knowledge about the individual structure, its components, and where the damage is located. If the damage level is, however, expressed only through different damage states, this tends to result in large uncertainties. In this case, the building response has to be defined in terms of probabilities for a building being in any of the damage states. Rather than a single set of fragility functions defining the probability for an intact building, a set of state-dependent fragility functions are required for each state. The framework proposed here therefore allows the assigning of a set of such fragilities to each building.

As a result of the time-variant hazard and repair process, the damage state is a variable feature of the building during the simulation and the fragility functions have to be adjusted according to the current damage state of a building. Within the proposed framework, the buildings go through the history of ground motions, are potentially damaged to some degree, may be repaired after some time, and then damaged more or again during subsequent events. The final result is a detailed damage and repair history for each of the considered buildings for a set of stochastic ground motions resulting from full seismic sequences.

In order to account for the probabilistic nature of the involved processes, a probabilistic Monte-Carlo simulation (Ulam et al. 1947; Metropolis and Ulam 1949) can be

employed to obtain a population of independent simulations. Associating the individual damage state transitions with an economic loss resulting from this transition allows a metric to be assigned, which makes the individual simulations comparable. Simulations can thus complement each other, but simulations following different modeling assumptions can also be compared. Therefore, the influence of various model components, such as Poissonian hazard compared to full seismic sequences, different time-delayed repair models, fragility changes with sustained damage, or the observation period, can be examined. Separating the individual processes into individual layers, i.e., stochastic generation of seismicity, calculation of ground motion, evaluation of fragility models, time-dependent repair processes, and loss estimation, makes the framework modular, allowing individual components to be replaced if better models are available or developed.

To summarize the points above, the innovative framework will employ a simulation-based approach that allows one to:

- model full seismic sequences over various observation periods,
- take spatial correlation of resulting ground motions into account,
- consider a large set of individual and spatially distributed buildings,
- respect the damage-state dependency of fragility,
- include either fixed-probability or time-dependent repair models.

A modular design of the proposed framework will furthermore facilitate a detailed sensitivity analysis of individual components.

3. Poissonian seismic risk on urban scale

Time- and state-dependent seismic risk assessment is an extension to this classical risk formulation. This chapter will thus introduce the concepts and components required to determine classical seismic risk on which the time and state dependent formulation in Chapter 4 will be based.

3.1. Poissonian seismic hazard

According to the original formulation of PSHA (Cornell 1968) the ground motion hazard at a site is described as the probability of exceeding a certain level of ground motion conditional on the occurrence of a seismic event with a certain magnitude within a certain distance during a defined time period. As outlined in the introduction, for the case of main shock (Poissonian) PSHA, the events are considered independent random events of a Poisson process. The probability of occurrence can thus be described as the probability of observing at least one event ($k > 0$) within a period t that causes a ground motion Y that exceeds ground motion level y :

$$P(k \geq 1|t) = 1 - e^{-\lambda t} \quad (3.1)$$

with an average rate λ obtained as:

$$\lambda(Y > y) = \sum_i^n \nu_i(M_i) \int_{r_{min}}^{r_{max}} \int_{m_{min}}^{m_{max}} f_{M;i}(m) f_{R;i}(r) P(Y > y|m, r) dr dm \quad (3.2)$$

for a level of ground motion y , where n is the number of considered sources, ν_i is the rate of earthquakes with a magnitude M_i at the source i with a magnitude distribution of $f_{M;i}$ (bounded by m_{min} and m_{max}), a source to site distance distribution of $f_{R;i}$ (bounded by r_{min} and r_{max}), causing ground motion exceedance with a probability of $P(Y > y|m, r)$. In the following sections, the derivation of these components is described.

3.1.1. Seismogenic sources

The first step in any seismic hazard assessment is the identification of the seismogenic sources able to cause shaking at the site of interest. As briefly discussed in

the introduction, earthquake ruptures usually occur on tectonic plate boundaries and pre-existing faults. Ideally, a precise record of past seismicity would delineate these structures sufficiently. In most regions though, these seismic catalogs are flawed in several ways. Instrumental recordings are usually available for only the past few decades. Depending on the density of the seismic networks, commonly increasing with time, many smaller magnitude events are missing. In the case of sparse networks, the location and magnitude estimates in these catalogs suffer large uncertainties. To extend the perspective to pre-instrumental times, these instrumental recordings can be complemented by the historical documentation of events. The quality of these historical reconstructions based on qualitative information collected by eye-witnesses or reporters of any kind strongly varies and the determined earthquake parameters can be considered as having a fairly high uncertainty. Additional data is thus usually required to identify potential earthquake sources. Such information includes geological surveys to identify faults, satellite-based analysis or deformation measurements from global navigation satellite systems. While catalogs usually only provide point locations, i.e., hypocenters of past earthquakes, the latter provide the actual delineation of the faults. A problem though is that these are usually limited to the delineation of faults close to the surface. Faults that reach the surface require a relatively large magnitude ($\sim M_w 6$), but smaller events can also be damaging. Furthermore, the fault structure at depth can only be investigated using geophysical methods, such as, e.g., seismics or seismic tomography. Very few faults are thus well mapped. To account for the resulting spatial uncertainties associated with the earthquake occurrence, where faults are not known, it is common to define spatially distributed seismicity. This can be achieved through, e.g., area source zones to either fully replace or complement fault information. Figure 3.1 gives examples of simple point sources, area source zones, and more complex sources defining the actual geometry of fault planes. Another popular, often complementary, method to consider spatial uncertainty is to use kernel windows to smooth the location of past seismicity to determine a spatial probability function, usually defined on a grid (Frankel 1995; Woo 1996). In most PSHA studies, alternative source models are constructed and combined in a logic tree, which is intended to capture the epistemic uncertainty arising from the lack of knowledge. Combining the alternative models, each contributes to the hazard estimate with a certain probability, which is usually assigned based on expert judgment, ideally reflecting the model's likelihood.

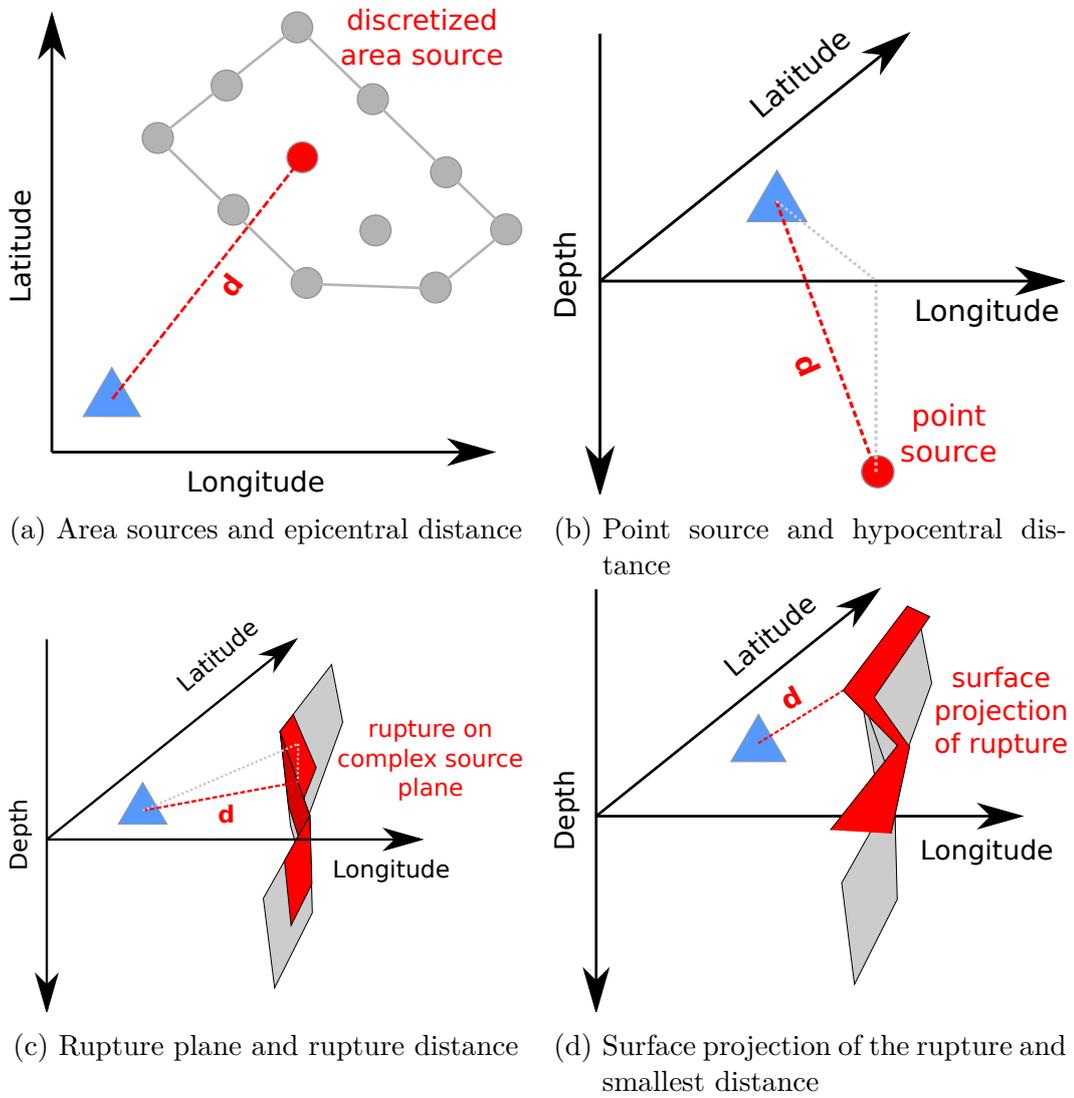


Figure 3.1.: Different approaches to modeling a seismic source (grey points and planes) as 3.1a area source zone, 3.1b point or 3.1c + 3.1d complex fault geometry. Each of the different source models allows different source-to-site distances d to be calculated, which may consider the distance from the site (blue) to the actual rupture geometry (red planes), or just a point in space (red points).

3.1.2. Source-to-site distance

The definition of seismogenic sources directly affects the source to site distance distribution (f_R in equation 3.2). If the complex fault geometry is well defined it can be used in the source model and the distances to the fault or its projection to the Earth's surface can be considered. The same is true for simplified fault planes, commonly described by linear segments on the horizontal plane, dipping angle, and nucleation depths. If physical faults are not defined, but only origin points over a set of points (area or grid) and no depth is provided, only the epicentral distance can be considered. However, if depth is provided, then more precise hypocentral distances can be calculated. If more sophisticated distance measures are to be considered, the distribution of potential rupture parameters must be defined in order to generate artificial ruptures around the nucleation points. Figure 3.1 illustrates some of these distance measures. The minimum required set of information for modeling an earthquake rupture as a plane in space are the nucleation point at depth, the strike, dip and rake angles of the rupture. The extent of the plane can then be estimated from magnitude scaling laws, such as those given by Wells and Coppersmith (1994). Modeling distributed seismicity rather than actual faults of course will usually lead to larger uncertainties and or spatial smoothing in the resulting hazard estimates.

3.1.3. Magnitude distribution and rate of occurrence

Once the set of seismogenic sources and their representation in the model is defined, the second step is to associate these sources with a rate with which these produce earthquakes and their magnitudes. In order to obtain seismicity rates for main-shocks (i.e., individual clusters) the seismic catalog is declustered, i.e., fore- and aftershocks are removed from the catalog to ensure only one event (the main-shock) remains per cluster. In order to decluster the catalog, methods such as Gardner and Knopoff (1974) can be used to determine earthquake sequences using space and time-windows (see, e.g., van Stiphout et al. 2012) and remove all, but the largest event in these clusters from the catalog.

A simple statistical relation for the frequency-magnitude relation was found by Gutenberg and Richter (1944), referred to as the cumulative Gutenberg-Richter law:

$$\log_{10}(N_M) = a - bM \quad (3.3)$$

where N_m is the number of observed events greater than or equal to a magnitude M within a certain period, a the total rate, and b is the ratio between large and small magnitudes. The constants a and b are commonly estimated from observed seismicity for each of the defined seismogenic sources. This regression analysis is sensitive to missing events in the catalogs and wrongly estimated rates may impact upon the hazard estimates derived from these rates. Figure 3.2 shows an example for observed cumulative counts of earthquakes and an estimated Gutenberg Richter law from Haas et al. (2016b). A common extension of the Gutenberg-Richter law

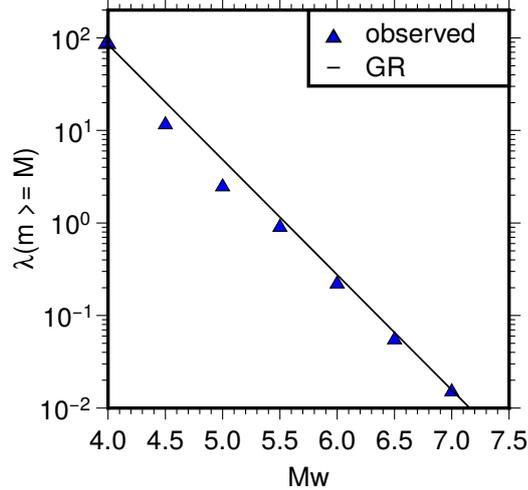


Figure 3.2.: Example for observed seismicity and the corresponding maximum likelihood Gutenberg-Richter relation estimated in Haas et al. (2016b, $a = 6.92$, $b = 1.25$).

is to employ a double truncated version with a minimum (m_{min}) and maximum magnitude (m_{max}), rewriting equation 3.3 as:

$$N_M = e^{\alpha - \beta M}, \quad (3.4)$$

with $\alpha = \log_{10}(e)a$ and $\beta = \log_{10}(e)b$, the magnitude probability distribution in the PSHA formulation (equation 3.2) can then be described as (Kramer 1996, p. 124):

$$f_M = \frac{\beta e^{-\beta(M - m_{min})}}{1 - e^{-\beta(m_{max} - m_{min})}} \quad (3.5)$$

The minimum magnitude is used to exclude events which are unlikely to produce damaging ground motion and the maximum magnitude is intended to avoid considering physically impossible ruptures due to the finiteness of the present faults. The latter is often determined from observed seismicity. The maximum magnitude and its derivation, however, are strongly disputed in the research community (e.g., Holschneider et al. 2011; Zöller et al. 2014; Zöller and Holschneider 2016). Assuming time-invariance in the tectonic process, the b -value for natural seismicity is a constant and can be determined using the most complete part of a seismic catalog. Complete means that from the magnitude of completeness on, the slope of the observed cumulative magnitude counts is not changing, i.e., the Gutenberg-Richter law holds. The magnitude of completeness (m_c) can be estimated for a catalog employing the median-based analysis of the segment slope (MBASS; Amorese 2007), which detects deviations of the median slope of the Gutenberg-Richter law and determines the magnitude of completeness if these deviations are significant. Commonly, the magnitude of completeness gets smaller over time as networks are improved and smaller magnitude events are detected. To assess the variation in completeness over time, catalogs are split into time intervals and the completeness in terms of (binned) magnitude is determined using, e.g., an algorithm such as that described by Stepp (1971) or a visual approach (see, e.g., Burkhard and Grünthal 2009), plotting cumulative number of earthquakes binned by magnitude (e.g., 0.5 M_w steps) against

years and for each magnitude bin identifying the year of completeness as the year where its rate changes significantly.

The Gutenberg-Richter b -value can then be estimated applying a maximum likelihood estimator (Aki 1965; Utsu 1965) using the average magnitude \overline{M} of events in the catalog larger than M_c :

$$b = \frac{\log_{10}(e)}{\overline{M} - M_c} \quad (3.6)$$

In addition, one can consider the uncertainties in the locations and magnitudes listed in the catalogs, defining probability distributions for both and employing a Monte-Carlo simulation (Ulam et al. 1947; Metropolis and Ulam 1949) to generate stochastic realizations of the catalog and determine a distribution for the b -value rather than a single value. Theoretically, the b -value should be close to one for natural seismicity.

In the simplest case, the two bounding magnitudes and the Gutenberg-Richter b -value are set to single values for all seismic sources in the region of interest. However, this is questionable, and especially the latter may result in a bias of the hazard. In any case, variations of the total activity, i.e., the Gutenberg-Richter a -value, have to be considered. As a depends on the deformation rates which are rarely constant over extended regions, it should be estimated for each source individually. In case this is not possible, e.g., because of a lack of data, a simple tool to obtain a spatial distribution of seismicity rates are the previously mentioned smoothed seismicity approaches. The approach proposed by Frankel (1995) spans a equally spaced grid over the region of interest. Then observed seismicity is counted within each grid cell, excluding events that are below the magnitude of completeness for the year of the event. This count n is corrected to n^* in each cell using the theoretical count of events expected from the Gutenberg-Richter law and the completeness estimates:

$$n^* = n \frac{\sum_i e^{-\beta m_i}}{\sum_i t_i e^{-\beta m_i}} \quad (3.7)$$

where m_i is the upper end of the magnitude bin and t_i the period for which this magnitude is complete in the catalog.

This count relates to the cumulative a -value as $n^* = 10^a$ and can be converted to an incremental a -value (a_{inc}) using the Herrmann (1977) formula, which relates them as:

$$10^{a_{inc}} = 10^a (10^{b\Delta M/2} - 10^{-b\Delta M/2}) \quad (3.8)$$

where ΔM is the magnitude bin width. To account for spatial variability, the method smooths the obtained counts over neighboring grid cells using a Gaussian kernel (cut off at 3σ). Figure 3.3 shows an example of gridded 10^a values calculated from the seismic catalog presented in the figure.

Using the Frankel (1995) approach on a declustered catalog yields a spatial probability as well as an average main-shock (or equivalent - sequence) rate model. If only a spatially smoothed seismicity model is used, the resulting seismogenic source model is particularly vulnerable to the incompleteness of the catalog, as seismic

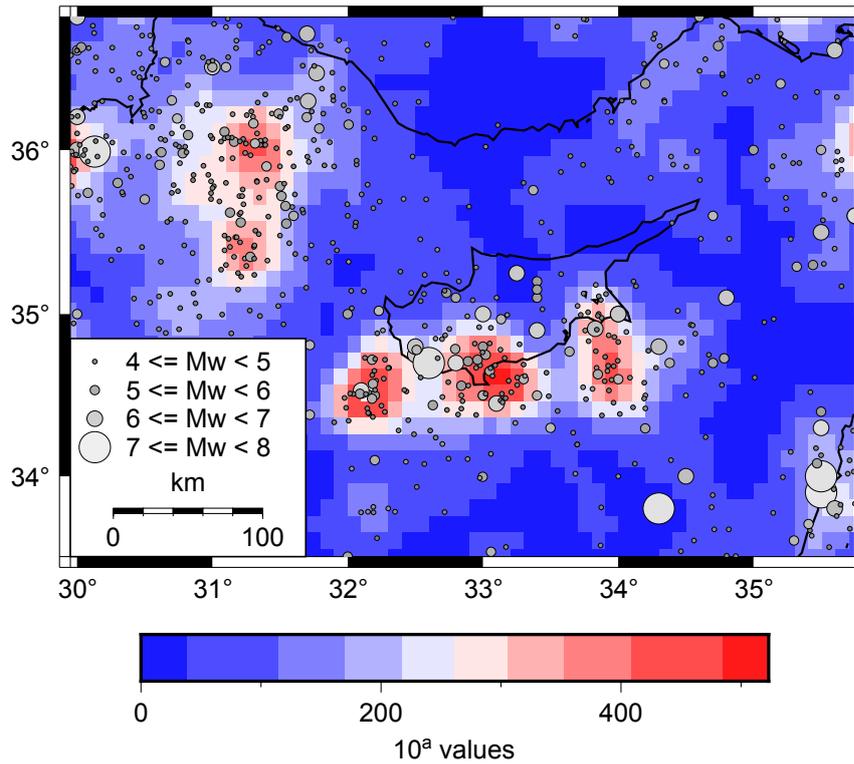


Figure 3.3.: Gridded Gutenberg-Richter 10^a values obtained using Frankel (1995) on the presented catalog (gray points) and employing a 0.1 by 0.1 degrees grid, a smoothing kernel with 20 km , and a minimum magnitude $M_w = 4$.

features which are not evident from the catalog might be active despite the lack of recordings.

3.1.4. Ground motion and site effects

As discussed in the introduction, the risk resulting from damage to structures depends on the intensity of the ground motion that a structure experiences. Having defined the probability model for the seismicity, the ground motion resulting from these earthquakes at the site of interest remains to be determined. Commonly, the intensity is described through a single parameter intended to characterize the actual ground motion. Such parameters can be peak ground displacement (PGD), peak ground velocity (PGV), peak ground acceleration (PGA), or spectral acceleration (SA), i.e., the response of harmonic oscillators with certain fundamental frequencies to the ground motion. Since in many regions, especially those prone to low or moderate levels of seismic hazard, instrumental records are still sparse, macroseismic intensity measures are still prominent. These are based on the earthquake's effects, which can be estimated from qualitative information like damage reports and the population's observation. Despite increasing instrumentation, this is still a commonly collected form of information, currently also gathered through web-based questionnaires (Atkinson and Wald 2007). Numerical, physics-based simulations of ground motion are gaining popularity (e.g., Bielak et al. 2010; Chaljub et al. 2010; Paolucci et al. 2014; Smerzini and Pitilakis 2018) as they are able to produce detailed ground shaking estimates. However, as these are computationally costly, and require detailed information on the rupture kinematics, travel paths, and site conditions such as the subsurface wave velocity structure, the use of regression models derived from observed data is still the most actionable solution for PSHA in many cases. These regression models describe the distribution of ground motion considering a limited set of parameters and are determined by fitting coefficients c_i of ground motion models in a form similar to (Kramer 1996, p. 87):

$$\ln(Y) = c_1 + c_2M + c_3M^{c_4} + c_5\ln(R + c_6e^{c_7M}) + c_8R + f_{source} + f_{site} + \sigma \quad (3.9)$$

to empirically observed amplitudes Y . These models consider different parameters such as ground motion scaling with magnitude (M), geometric spreading with distance (R), finiteness of the rupture (depending on M), inelastic attenuation as the wave passes through the rock (depending on R) and properties regarding the source and site. Modern models such as Campbell and Bozorgnia (2014) have 41 different coefficients and parameters commonly estimated for 20 or more different periods for the elastic response spectra. As these are based on statistics, they have an inherent uncertainty, and as indicated in equation 3.9, this is usually described by a log-normal distribution. From this distribution, the exceedance probability as defined in the PSHA formulation ($P(Y > y|m; r)$ in equation 3.2), conditional on the event, path and site parameters, can be determined (note that in equation 3.2 only magnitude and distance are indicated). The residual distribution defining the aleatory uncertainty can be split into a inter- and intra-event residual (Abrahamson and Youngs 1992). This means that the ground motion varies from event to event, but also the very same event might cause different levels of ground motions at the same site. This uncertainty is also attributed to the need to combine observations from different sites, with to some extent similar conditions, since for many sites, there are no sufficient numbers of recordings. With the increasing availability of recordings in recent years, site-specific GMPEs are available for an increasing number of sites (e.g., Rodriguez-Marek et al. 2011; Abrahamson et al. 2014; Kotha

et al. 2016). However in most cases, such models are not available and this leads to increased epistemic uncertainty, as it is unclear how the ground will respond at a site. Within PSHA, this uncertainty is considered through logic trees, where alternative GMPEs are defined as the tree’s branches. In order to capture the epistemic uncertainty realistically, the selected attenuation models should best describe the observed ground motion attenuation. At the same time, redundancy must be avoided in order to not bias the result. A recent collection of globally available GMPEs by Douglas (2017) contains 432 GMPEs for PGA, 278 for response spectra estimates, and 41 macroseismic intensity attenuation models. In order to choose which of these models are suitable for implementation in the logic tree, analysts commonly apply different methods using available ground motions records for the region of interest. These methods are often either based on a visual comparison of models and data, or they employ different statistical ranking methods proposed in literature (e.g., Scherbaum et al. 2009; Kale and Akkar 2013).

As has been observed from many past events, the local soil conditions, especially in basins with loose sediments, can lead to large amplification of the ground motion with respect to sites on solid rock (see, e.g. Ansal 2004). In general, soil modulates ground motion amplitudes over various frequencies and changes the duration of shaking. These effects can be modeled analytically or empirically. In most cases, GMPEs also contain a site term (see equation 3.9), this term uses independent variables which correlate with ground motion. Common variables are the shear wave velocity in the upper 30 meters ($v_{s,30}$), or the depths at which the shear wave velocity drops to 2.5 and 1 km/s , respectively. While the former shows some correlation with surface effects, the latter two are intended to capture deep sedimentary basin effects. This is often not sufficient to capture the effects realistically, but in many cases it is used as a first-order approximation.

The previous descriptions define the distribution of ground motion at a single site. In case of an urban seismic risk assessment however, the ground motion is not estimated at a single site, but at a number. The ground motion caused by an event shows spatial correlation between these sites, meaning that if, e.g., at a given site a ground motion exceedance of the median value occurs, neighboring sites are likely to exceed this value and vice versa. This can be modeled using an isotropic spatial correlation model as defined by Jayaram and Baker (2009).

3.2. Fragility and vulnerability of buildings

From the building perspective, the hazard related to ground motions is mainly due to lateral forces and moments induced as the motion are transferred to the structure. If the resulting deformations in the building elements exceed their elasticity, these are damaged. In large scale probabilistic risk analysis, the complex damage process experienced by a structure is commonly summarized by discrete levels of damage such as *slight* (ds_1), *moderate* (ds_2), *extensive* (ds_3), and *complete* (ds_4 ; HAZUS; Federal Emergency Management Agency 2011). These limit states are defined through an observable engineering demand parameter, in the simplest case

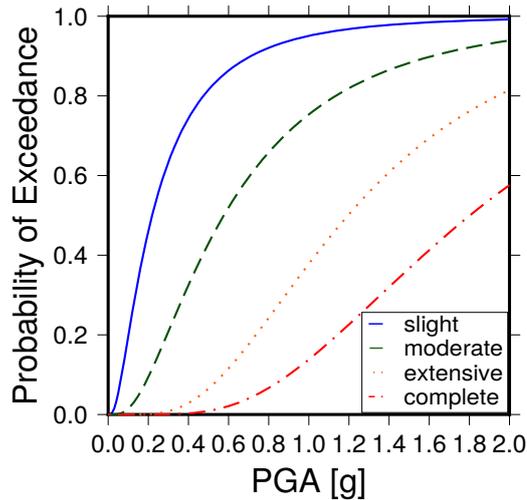


Figure 3.4.: Fragility curves describing the probability of exceeding a building’s response corresponding to different damage states (here HAZUS).

displacement, e.g., roof displacement d as in (Lagomarsino and Giovinazzi 2006):

$$\begin{aligned}
ds_1 &= 0.7d_y, \\
ds_2 &= 1.5d_y, \\
ds_3 &= 0.5(d_y + d_u), \\
ds_4 &= d_u.
\end{aligned}
\tag{3.10}$$

where d_y and d_u are yielding and ultimate levels of roof displacement at which a building starts to deform permanently or is expected to fail, respectively. Fragility of a structure is then the probability of exceeding a certain level of damage conditional on experiencing a given level of ground motion.

Usually, fragilities are described with exceedance probabilities $1 - F(x)$ of these levels ($x = X$), using lognormal distributions, with a cumulative distribution function (CDF) expressed as:

$$F(x) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\ln(x) - \mu}{\sigma\sqrt{2}} \right) \right]
\tag{3.11}$$

with a mean μ and standard deviation σ .

There are several methods to derive fragility functions for structures, with varying complexity and associated analysis costs. The two extremes are the cost extensive assignment of existing fragility models through visual screening and the cost intensive modeling of structures using sophisticated finite element models (Vamvatsikos and Pantazopoulou 2016). While the insight in expected seismic behavior of the former is limited, the latter requires very detailed information on various parameters of the structure. Additionally, as pointed out previously, building response changes significantly if a structure is damaged (Aschheim and Black 1999). This behavior is of course complex to model as already small changes in the structure can result in strong variations in its response. While feasible on the single building scale, the cost of analysis for dynamic finite element analysis over urban scales are usually not justified, considering the large uncertainties involved (Vamvatsikos and Pantazopoulou 2016). Thus over urban scales, rather than determining building response

for each individual structure, buildings are grouped into building types which are likely to show similar seismic behavior. This is to simplify the assessment, ideally maintaining accuracy on aggregated, e.g., urban or quarter, scales, at the cost of loosing precision and often also accuracy at the single building scale.

For the case of direct economic loss assessments, damages are translated to fractions of the costs resulting from a theoretical complete loss of the structure, referred to as loss ratios. They are often considered as a fixed percentage (Federal Emergency Management Agency 2011), although in reality these vary from case to case. One can also obtain a vulnerability function for a structure by summing over loss ratios multiplied by the probabilities of the corresponding damage states for given levels of ground motion as described by the fragility function.

3.3. Exposure

If a seismic risk assessment is conducted at urban scales, identifying the assets exposed to the hazard is a crucial component of the analysis. On large scales, exposure is a highly dynamic and uncertain component of risk (Pittore et al. 2017). For simplicity, exposure is often reduced to buildings. However, as outlined in the previous section, from a fragility perspective each building can be considered an individual. Many of their features vary, and as a result so does their structural response, e.g., to seismic loading. The aim of the exposure analysis is to determine the spatial and featural distribution of buildings as detailed as possible, in order to constrain the building stock's fragility. There are two potentially complementary approaches to gather exposure information: desktop studies and field surveys. While the former collects (often aggregated) information available from governmental institutions, ancillary data, or previously conducted studies, the latter collects information on individual buildings in the field. Due to the strongly dynamic nature of exposure (Pittore et al. 2017), both cases yield only a snapshot of the exposure at the moment of data collection.

Some countries assess their buildings within the scope of a national census. These housing census data are usually publicly available only on highly aggregated scales, such as national, district or sub-district levels. The focus of the census data collection is commonly more on parameters dealing with the quality of life or human development, and only a few structurally relevant building features are collected. In many cases, these are limited to wall materials which do not always allow building typologies to be unambiguously associated, e.g., masonry walls can be found in masonry, but also infilled frame concrete buildings. As the building data collection is usually not conducted by engineers, this information is also relatively uncertain. Census data can be complemented by ancillary data such as building footprints or road networks from crowd sourced data, for example OpenStreetMap (OSM; OpenStreetMap contributors 2017). Despite the limited level of detail on relevant structural information and only a loose connection to fragility, these data can be used to build a preliminary large scale exposure model. The derived model will be a statistical model subject to large uncertainties (Haas et al. 2016d). If the extent

of the considered agglomeration is large enough to show strong variations in ground motions, this data has to be spatially deaggregated. A common approximation is usually obtained using spatial population density estimates. Such an approach, i.e., deaggregating high level data to derive features for low level entities is referred to as top-down approach (Fekete et al. 2010).

Such an exposure model can be improved if more detailed data on individual buildings is collected in the field. Ideally, this allows better constraints on the structural components of the buildings and their spatio-geographical distribution. However, a full building census is only feasible for very small agglomerations, hence in most cases the assessment is limited to a sampling. In order to represent the population, the sampled data is aggregated afterwards, which is referred to as bottom-up approach. The resulting exposure model will also be a statistical model, however, as the level of detail on the individual entities of the population is higher, the latter should yield a better model than desktop studies (Wieland et al. 2012; Pittore and Wieland 2013).

3.3.1. Exposure data sampling

Collecting representative samples of a population requires a thorough sampling approach. An effective way to gain information about a heterogeneous population is to divide it into homogeneous strata and sample these from these. Considering the size of the strata more reliable statistics on the whole population can be deduced than if sampling the whole population randomly. For this purpose, a cost efficient stratified sampling approach based on multi-source imaging was proposed recently (Wieland et al. 2012; Pittore and Wieland 2013; Wieland et al. 2015). It comprises a multi-tier analysis framework (see Figure 3.5) which involves the use of imagery with increasing resolution and level of detail of gainable information. Tier 1 is based

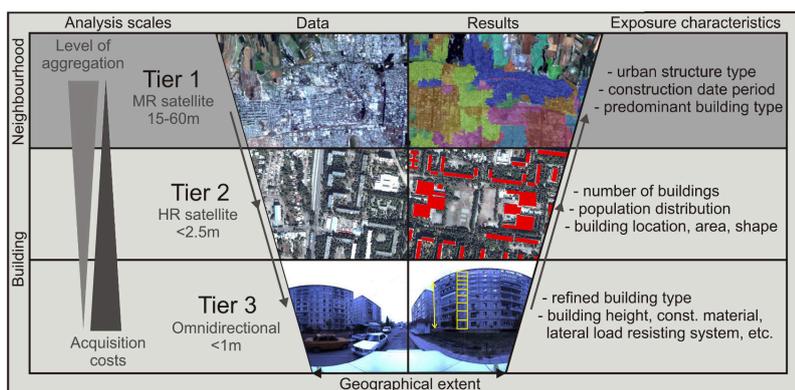


Figure 3.5.: Multi-tier imaging analysis for gathering exposure data, with analysis scale, example images, and obtainable information (modified after Wieland 2013).

on multi-spectral satellite imagery analysis with medium spatial resolution. These images are not sufficient to reliably identify individual structures whose dimensions are commonly below the image resolution, but can be used in order to subdivide a

town into regions (strata) with a (relatively) homogeneous built environment. The proposed approach employs object based imaging using machine learning techniques and relies on well defined ground truth data for training and testing purposes. In many cases available ground truth data is limited, for this case a simplified pixel based procedure might be sufficient for a preliminary analysis (see Appendix B for a description). The finer granularities (Tier 2 and Tier 3) comprise methods to obtain information on the level of individual structures. In an imaging analysis framework, this can mean using high-resolution satellite imagery to obtain information about building density, as well as mainly geometric characteristics of individual buildings (Tier 2), however this imagery is relatively expensive.

While top view imagery can provide information on the geometry and spacing of structures, the derivation of further building attributes is limited. Thus, a complementary data source are ground based surveys (Tier 3). These yield the most detailed exposure information within this framework and can be obtained by rapid visual building screening of omni-directional imagery collected with a mobile mapping system. The path that is taken with the mobile mapping system is based on the satellite analysis of Tier 1 and aims to collect imagery for buildings in each of the strata. Routines and tools for collecting and processing the acquired imagery to obtain building data in a standardized way employing the GEM Building Taxonomy (Brzev et al. 2013) are described in Appendix B.

Assuming that Tier 1 satellite analysis can indeed yield a separation of a town into homogeneous regions, the approach can be used in a bottom-up approach, i.e., integrating low-level data in the strata obtained from the satellite analysis.

3.3.2. Building types

Individual building data is collected with the purpose of deriving or associating fragility models. The most common approach is the direct assignment of predefined building typologies, which in most cases are directly linked to a set of fragility functions. Such categorizations are proposed, e.g., in ATC-13 (Rojahn et al. 1985), the European Macroseismic Scale (Grünthal et al. 1998), FEMA 547 (Federal Emergency Management Agency and Chekene 2006), or Jaiswal and Wald (2008). The fundamental problem with this approach though is that the assignment of building categories might not always be fully reproducible, since the observations and the reasoning leading to the building type assignment are not necessarily documented. This has several drawbacks, first of which is that the assignment can be subjective in the sense that different engineers might assign different typologies. Building typologies are not easily modifiable, i.e., one cannot add additional typologies afterwards or split categories easily into subcategories if new information becomes available. With these problems in mind, and inspired by a taxonomy proposed in the SYNERG project (Crowley et al. 2011; Hancilar and Taucer 2013) that aimed to hierarchically classify buildings through their attributes, the Global Earthquake Model (GEM) initiated a change of paradigm by moving towards a purely observation-based exposure data collection. Rather than describing buildings by fitting them

into predefined categories, engineers collect individual building attributes using a standardized attribute-based taxonomy. This taxonomy, called the GEM building taxonomy (Brzev et al. 2013), defines 12 attribute classes describing building material and properties, lateral load resisting system, height, age, occupancy, position, shape, regularity, walls, features of the roof and floors, and foundations. In addition, the buildings' attributes can be described separately for two directions, e.g., parallel and perpendicular to the street. Each of the attribute classes contains one or more attributes, where most of these are standardized. In total, up to 24 different attributes describe an individual building. The application of the taxonomy yields for each building a description of its features in the form of a long collection of the standardized attributes (represented by short, unique codes), comprising what is called a taxonomy string. Such a taxonomy string for a low rise building with unreinforced masonry walls defined as URML in FEMA 457 with many attribute values set to unknown (strings with 99) could look like the following: MUR+MUN99+M099/LWAL+DU99/HBET:2,1+HF99/Y99/OC99/BP99/PLF99/IR99/EWMA/RSH99+RMT99+R99+RWC99/F99+FWC99/FOS99 (shortened from Brzev et al. 2013). This string could be seen as the DNA of a building.

As a result, the richness of a dataset collected with an attribute-based taxonomy such as the GEM building taxonomy, is unmatched by taxonomies that rely on predefined building types. With this data as the basis, any derived exposure model is easily refinable, expandable and reproducible. In addition, uncertainties in the observation can be more directly assessed as only direct observations are collected and uncertain attributes are defined as unknown. One might even collect degrees of belief for individual attributes, as proposed by Wieland et al. (2015). However, the application of the GEM building taxonomy comes at a cost, as the observations, differently to the direct assignment of building typologies, are not explicitly linked to a fragility model. This can be seen as a drawback as post-processing is more complex because fragility models have to be derived for the individual observed building configurations or links to existing fragility models have to be defined. On the other hand this allows for a flexible and modifiable fragility assignment.

The complexity of the post-processing becomes clear when looking at an example. On the scale of a town, a data set following the GEM building taxonomy can easily result in a large set of different attribute configurations. Even small towns can have potentially hundreds of unique building attribute combinations. Figure 3.6 shows a graph plot that gives an impression of the complexity of such a data set for about 800 buildings. Although the example here stems from a relatively homogeneous population, many unique building types (red squares) can be found. These unique building types are defined via their unique combination of attributes (blue circles, connected via gray lines to the building type). Although many attributes are shared among different building types, there are about 700 unique combinations among the 800 buildings. Ideally, since most features are influencing structural behavior, an engineer would derive fragility and vulnerability curves for each of these combinations. Practically, as indicated earlier, this large number of types has to be reduced to a number actionable within the framework of a large scale seismic risk analysis.

The difference with respect to traditional exposure data collection is that the stan-

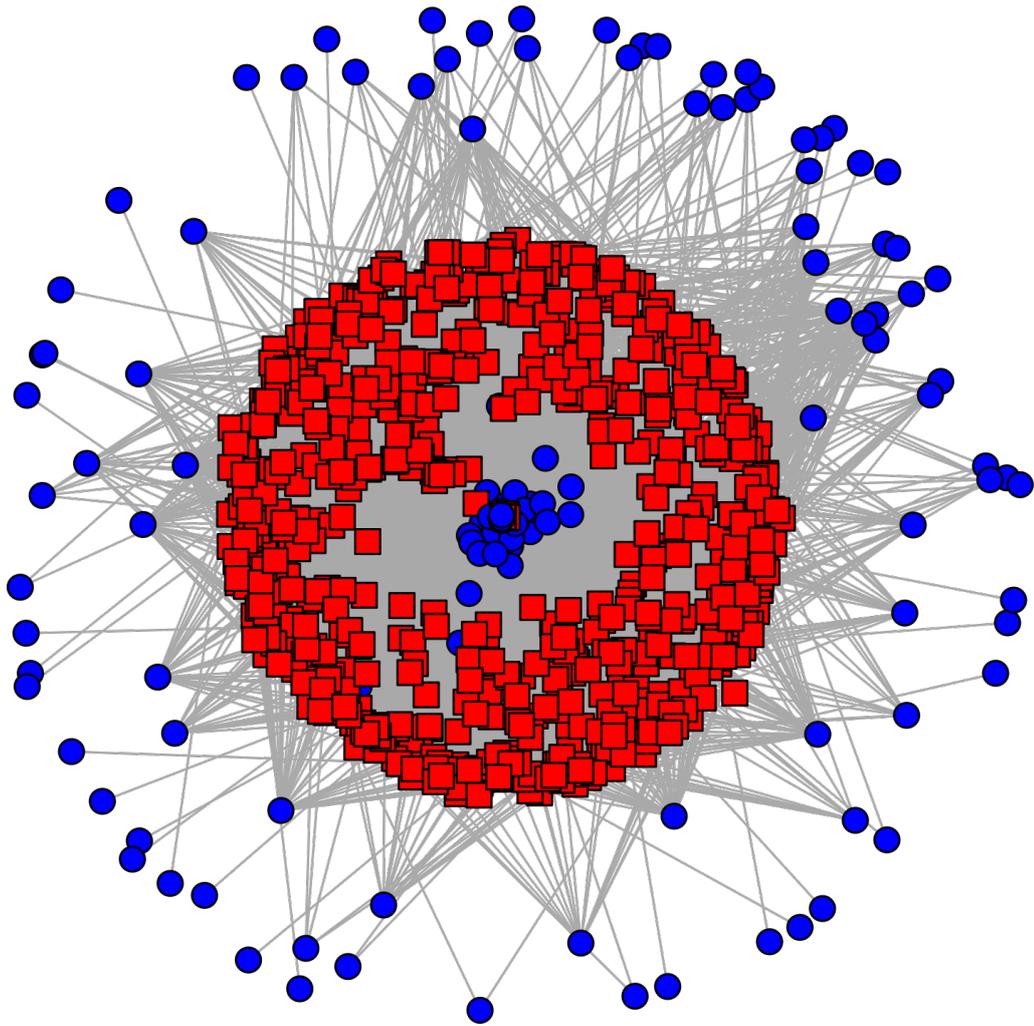


Figure 3.6.: Complex exposure data set collected with the GEM Building Taxonomy, where red squares are unique building configurations and blue circles correspond to all different encountered attributes which are connected to the building configuration through gray lines. Considering that the underlying dataset consist of only 800 buildings, the complexity, i.e., the unique set of building configurations is very large.

standardized description of the building features allows one to revisit and refine this reduction at a later point in case resources for a more thorough analysis become available. There are several ways to reduce the number of typologies. The final goal being to group the buildings with similar fragilities into more general types. A common approach is the definition of mapping relations or considering only a subset of the collected information (see, e.g, Yepes-Estrada et al. 2017). Given the level of detail, this can be a highly resource intensive task that, moreover, may be subjectively biased. The starting point for the definition of the mapping should thus be a thorough data analysis to identify similarity among the building types. A simple unsupervised clustering approach to this is described in Appendix B.7. A supervised approach for such mapping which employs the fuzzy scoring of few relevant attribute values as proposed in Pittore et al. (2018) is more costly, but can also explicitly account for the uncertainty in such mappings.

4. Considering time and state dependency

While the components described in Chapter 3 are sufficient to model seismic risk under classical assumptions, they have to be extended for time- and state-dependent seismic risk. The first required change is to consider full seismic sequences rather than only the mainshocks. As a result, the assumption that buildings are intact in subsequent events has to be revised, with a suitable repair mechanism and a more complex state-dependent description of damage needs to be adopted. The following sections describe these extensions and how they can be implemented into a risk simulation framework.

4.1. Modeling full seismic sequences

The spatio-temporal clustering of earthquake sequences can be modeled following the approach outlined in Beauval et al. (2006) and recently Moradpour et al. (2014) using an Epidemic-Type Aftershock Sequence (ETAS) model (Kagan and Knopoff 1987; Ogata 1988). This can be used to generate probabilistic realizations of full sequences of seismicity within a specified time period, starting from a Poissonian model as described above that drives the cascading process of natural seismicity. In the ETAS model, the total rate of earthquake occurrence with magnitude m at a point \vec{r} in space and at time t can be described as the overlay of a time-independent (Poissonian) background seismicity rate μ and a rate of events ϕ_{m_i} triggered by all events i (main-shock and triggered events) at point r_i and time t_i previous to t (Moradpour et al. 2014):

$$\lambda(t, \vec{r}, m) = \mu(\vec{r}) + \sum_{t_i < t} \phi_{m_i}(\vec{r} - \vec{r}_i, t - t_i) \quad (4.1)$$

Please note that triggered in the sense described here and throughout the remainder of this thesis refers to naturally triggered seismicity due to static and dynamic stress changes caused by ground motion, not human-induced events as might occur due to pore pressure changes or mass redistribution during or after mining, fluid extraction or injection, or hydro-fracturing activities. The rate $\mu(\vec{r})$ can be split into a regional constant μ and a normalized and discrete spatial probability model. Both can be obtained from annual Gutenberg-Richter a -values, determined using the previously outlined Frankel (1995) approach on a grid where for each grid cell (with index

i, j):

$$\begin{aligned} \mu(\vec{r}) &\sim \mu_{i,j} = \mu P_{i,j} \\ &= \frac{\sum_{i,j} a_{i,j}}{365.25} \frac{a_{i,j}}{\sum_{i,j} a_{i,j}} \end{aligned} \quad (4.2)$$

Recollecting that earthquakes show fractal behavior (Kagan and Knopoff 1987), each of the independent Poissonian events derived above triggers a sequence of events which themselves initiate further sequences. This is modeled in the ETAS model as a superposition of all triggered sequences, i.e., the sum in equation 4.1. Sequences can span several decades, thus one has to consider also that in addition to the background activity within a considered period, triggering events could also have happened previous to this period. There exist two strategies when considering this in a model. The first option is to use a sufficiently long forward simulation and use only the last years of the resulting catalog (see e.g., Papadopoulos and Bazzurro 2018). A second computationally simpler, but still realistic, choice is to simply add a historical (non-declustered) catalog prior to the observation period. This of course can result in an increase in the resulting hazard estimates if strong events were observed in the catalog, as demonstrated by Papadopoulos and Bazzurro (2018). Of course, the contrary may also occur in the case of lower than average seismic activity observed in the catalog.

Each of the catalog and background events trigger a sequence of aftershocks whose magnitudes show the same Gutenberg-Richter b-value as estimated for the background seismicity. The rate of triggered events ϕ_{m_i} (see equation 4.1) depends on the magnitude m_i of the triggering event, and can be described as a spatially inhomogeneous (depending on \vec{r}) and time-varying marked Poisson process (Moradpour et al. 2014):

$$\phi_{m_i}(\vec{r} - \vec{r}_i, t - t_i) = \rho(m_i) \psi(t - t_i) \zeta(\vec{r} - \vec{r}_i) \quad (4.3)$$

where $\rho(m_i)$ is the total number of events spawned by an event of magnitude m_i , with a normalized time $\psi(t - t_i)$ and spatial distribution $\zeta(\vec{r} - \vec{r}_i)$.

The total number $\rho(m_i)$ is following the power law dependence of perturbed stress volumes with the magnitude of the triggering event (Helmstetter and Sornette 2002):

$$\rho(m_i) = K 10^{\alpha(m_i - m_0)} \quad (4.4)$$

where K is the productivity of the sequence, m_0 is the lowest considered magnitude, which should be greater than or equal to the magnitude of completeness $m_0 \geq m_c$, and α is a regional parameter controlling the increase in the number of triggered shocks with the magnitude of the triggering event. For simplicity, only the case $\alpha = b$ is considered here, which corresponds to assuming that triggered events correlate with the rupture size.

The time distribution of the aftershocks can then be described by the modified Omori Law (Utsu and Ogata 1995, equation 2.1).

As $\psi(t - t_i)$ (equation 4.1) is the normalized time distribution of events at a time t after the triggering event at time t_i , equation 2.1 has to be divided by the time

integral over equation 2.1, corresponding to the total number $N(t)$ of aftershocks for a single sequence up to a point t . With $p > 1$ and t large enough, one obtains from this a rate $\psi(t - t_i)$:

$$\psi(t - t_i) = \frac{(p - 1)c^{p-1}}{(t - t_i + c)^p} \quad (4.5)$$

The inter-event time Δt can then be sampled from an inhomogeneous Poisson with the rate defined as above:

$$P(t \leq \Delta t) = 1 - e^{-\rho \int_{t_i}^{t_i + \Delta t} \psi(t) dt} \quad (4.6)$$

The last component of equation 4.3, the normalized spatial distribution $\zeta(\vec{r} - \vec{r}_i)$ of aftershocks around the triggering event considers the linearised density, i.e., the probability function of after shock distances, to vary as (Moradpour et al. 2014):

$$P_m(|\vec{r}| = r) = \int_0^{2\pi} \zeta_m(\vec{r}) d\phi \quad (4.7)$$

Based on empirical evidence, Gu et al. (2013) and Moradpour et al. (2014) describe the spatial density with two power laws: a close distance regime from half of the rupture length until 10 km (order of crustal thickness) and one beyond 10 km. A characteristic length scale stems from the rupture length that can be estimated with:

$$L_R = l_0 10^{\sigma m} \quad (4.8)$$

with constants l_0 usually around 10-20 m and σ around 0.4-0.5, which can be observed from the linear aftershock density with distance (Moradpour et al. 2014), which shows a maximum at around $L_m = L_R/2$. The density can then be described by (Moradpour et al. 2014):

$$P_m(r) = \frac{qr^\gamma}{L_m^{\gamma+1} \left(\frac{r^{\gamma+1}}{L_m^{\gamma+1}} + 1 \right)^{1 + \frac{q}{\gamma+1}}} \quad (4.9)$$

Considering the two distance regimes ($r \leq 10km$ and $r > 10km$) and normalizing to 1 using two factors α and β , the distance probability can be described with (Hainzl et al. 2014; Moradpour et al. 2014):

$$P_m(r) = \begin{cases} \alpha \frac{qr^\gamma}{L_m^{\gamma+1} \left(\frac{r^{\gamma+1}}{L_m^{\gamma+1}} + 1 \right)^{1 + \frac{q}{\gamma+1}}} & ; r \leq 10km \\ \beta \frac{dr^\gamma}{L_m^{\gamma+1} \left(\frac{r^{\gamma+1}}{L_m^{\gamma+1}} + 1 \right)^{1 + \frac{d}{\gamma+1}}} & ; r > 10km \end{cases} \quad (4.10)$$

where d is a constant similar to q describing the decay for distances greater than 10 km . The normalizing constants are (Hainzl et al. 2014):

$$\alpha = \frac{r^{\frac{q}{1+\gamma}}}{r^{\frac{q}{1+\gamma}} + \frac{q}{d} - 1}; \beta = \frac{q}{d} \frac{r^{\frac{d}{1+\gamma}}}{r^{\frac{q}{1+\gamma}} + \frac{q}{d} - 1} \quad (4.11)$$

The direction of the aftershock from the main shock is chosen randomly. In order to estimate the productivity K , Båth's law 1964 can be used, which expects an average magnitude difference between the main- and largest after-shock of $\Delta m = 1.2$. Deriving the number of aftershocks, considering that the productivity scales with magnitude as $10^{\alpha(m_i - m_0)}$ (equation 4.4) and the time integral of the Omori law, gives:

$$N_a = K 10^{\alpha(m_i - m_0)} \frac{c^{1-p}}{p-1} \quad (4.12)$$

According to the Gutenberg-Richter law, the number of aftershocks greater than or equal to m is determined as:

$$N(M \geq m) = N_a 10^{-b(m - m_0)} \quad (4.13)$$

Now, considering Båth's law, the second largest event is on average $m_i - \Delta m$, this means that:

$$N(M \leq m_i - \Delta m) \stackrel{!}{=} 1 \quad (4.14)$$

and thus:

$$1 = N_a 10^{-b(m_i - \Delta m - m_0)} \quad (4.15)$$

With the special case $\alpha = b$, this yields:

$$K = 10^{-b\Delta m} \frac{p-1}{c^{1-p}} \quad (4.16)$$

Repeatedly simulating background and triggered events results in a Monte Carlo framework yields a set of stochastic realizations of seismicity for the period of observation. If the ETAS is applied horizontally, depth and rupture information have to be modeled separately, e.g., through a simple spatial distribution.

4.2. State dependent damage and time dependent repair

If the time in between events is not sufficient to fully repair the damaged structures, a time-dependent probabilistic repair model has to be introduced. Using exponential models, Yeo and Cornell (2009b) suggested relatively short repair times of between 6 and 360 days, while Cesare et al. (1992) and recently Iervolino et al. (2016) suggest fixed repair probabilities for units of time. In reality the decision to repair a structure depends on factors related to the individual owners, but also external factors. The former include factors such as the survival of the owners, their financial situation, if the property was insured, the degree of damage and other subjective motivations. External factors include applicable laws, governmental support, the total number of damaged structures, and availability of construction personnel and resources.

Considering the involved complexity resorting to a statistical model seems reasonable. Figure 4.1 shows a global data set on reconstruction with time after a seriously

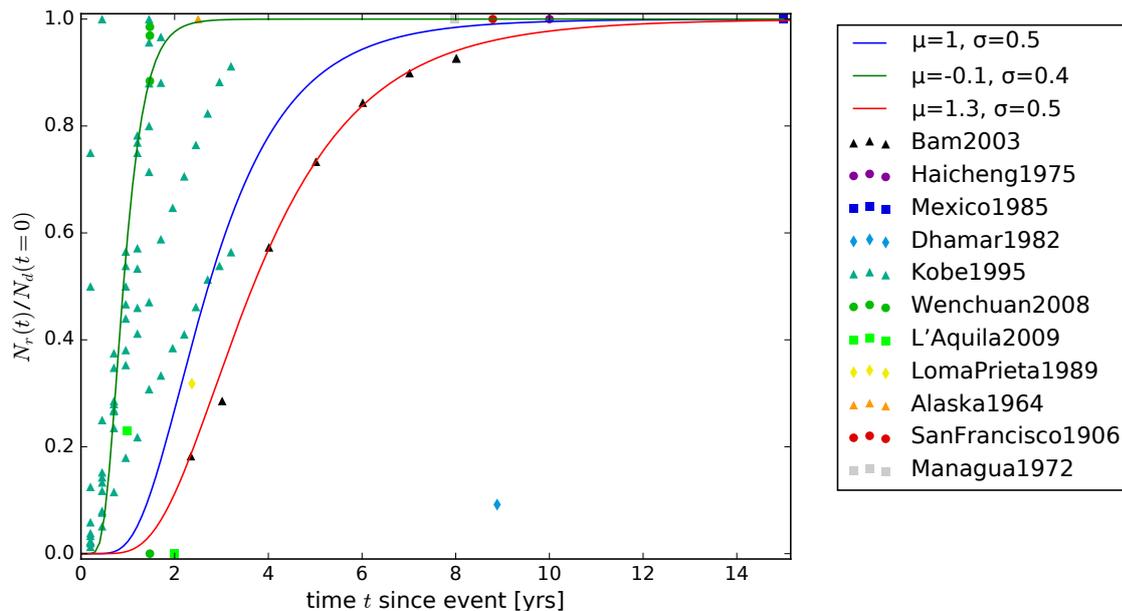


Figure 4.1.: Observed reconstruction ratio with time in years for several globally observed large events and three time dependent lognormal models tested within this thesis: optimistic ($\ln(\mu) = -0.1$, $\ln(\sigma) = 0.4$ [yrs]), conservative ($\ln(\mu) = 1$, $\ln(\sigma) = 0.5$ [yrs]), and pessimistic ($\ln(\mu) = 1.3$, $\ln(\sigma) = 0.5$ [yrs]).

damaging earthquake event collected within this thesis (Barakat 2003; Cheng and Wang 1996; Dunford and Li 2011; Dynes and Quarantelli 1989; Haas et al. 1977; Hashemi-Parast et al. 2017; Hirayama 2000; Liel et al. 2013; Russell 1994). The x-axis represents time in years passed since the event and the y-axis is the number of repaired or reconstructed buildings within a town ($N_r(t)$) with respect to the ratio of buildings damaged during the event ($N_d(t=0)$). A first observation from this plot is that the rate of reconstruction varies greatly. The time until full recovery in terms of building numbers ranges from less than a year until almost 15 years. This stresses that, a fixed probability model is inadequate on long time scales, and the short time periods assumed by Yeo and Cornell (2005) are probably too optimistic over large scales.

The plot is dominated by a relatively large data set for the case of the 1995 Kobe event (Hirayama 2000). Figure 4.2 shows a plot for the Kobe 1995 data, for whole Kobe (grey circles) and per affected quarter. The trends show reconstruction was carried out with relatively constant rates, although strongly varying for the different districts, showing only a slight saturation after the first till second year for the quarters of Nagata and Suma. While the quarters that lost only few of their buildings (Kita 2%, Nishi 2%) recovered relatively quickly, the strongly affected Nagata (39% destroyed) recovered the slowest, while the quarters close to the average rate show less strong correlations to percentage range of buildings lost from 4 until 25%.

The only other earthquake for which such a well-sampled time series is available from the data collected within this thesis is the Bam 2003 earthquake (Hashemi-Parast et al. 2017, see Figure 4.1). Unlike the Kobe case it shows a trend following almost

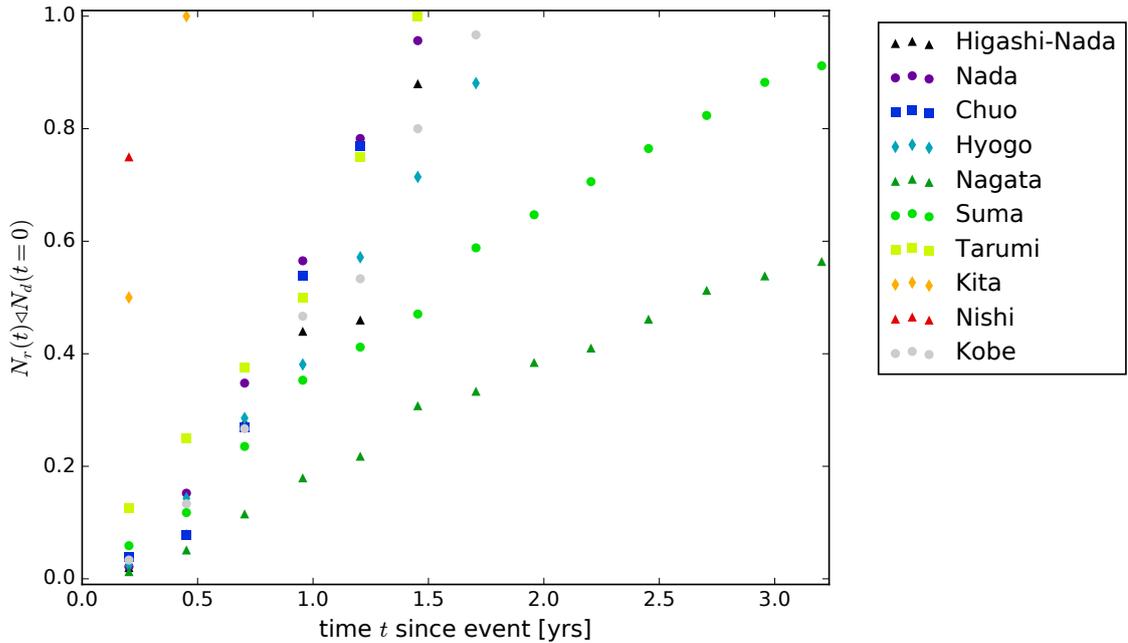


Figure 4.2.: Observed reconstruction ratios for Kobe (Japan) and its quarters affected by the 1995 earthquake (data from Hirayama 2000); gray circles are the average rate for the whole town.

perfectly a cumulative lognormal distribution. A peculiar case is the 2009 L'Aquila event (Liel et al. 2013), which shows two very different observations highlighting how the reconstruction of the central city is much behind that of the surroundings (in 2011). All other data points are scattered in between the Kobe and Bam observations, with the exception of Dhamar (Yemen), which even 9 years after the event shows a reconstruction ratio of less than 10%.

The figures above do not distinguish per damage grade as among the examined references such data is only available for a single time interval after the Bam 2003 event, for which Hashemi-Parast et al. (2017) find a strong increase with the damage grade of the ratio of repaired buildings with respect to non-repaired buildings in the time between 2003 and 2012 ($dg_1 : 0.59$, $dg_2 : 1.23$, $dg_3 : 2.86$, $dg_4 : 6.9$, and $dg_5 : 36.5$). Concurrently, they found also an positive correlation of buildings left vacant with increasing damage grade ($dg_1 : 0\%$, $dg_2 : 2\%$, $dg_3 : 19\%$, $dg_4 : 13\%$, and $dg_5 : 25\%$). To what extent this behavior can be generalized is currently unclear.

As a consequence of the previous observations, this thesis will distinguish between two types of probabilistic repair models. The first is a simple time-dependent model of lognormal shape:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln(t)-\mu)^2}{2\sigma^2}} \quad (4.17)$$

where $t > 0$ is the time since the damaging event, μ and σ the mean and standard deviation, respectively. The second option is a repair model with fixed rate, as suggested by Cesare et al. (1992) and Iervolino et al. (2016). To account for the strong variations observed in the state of repairs, three different lognormal time-dependent models are proposed. These are a *optimistic* model ($\ln(\mu) = -0.1$, $\ln(\sigma) = 0.4$ [yrs])

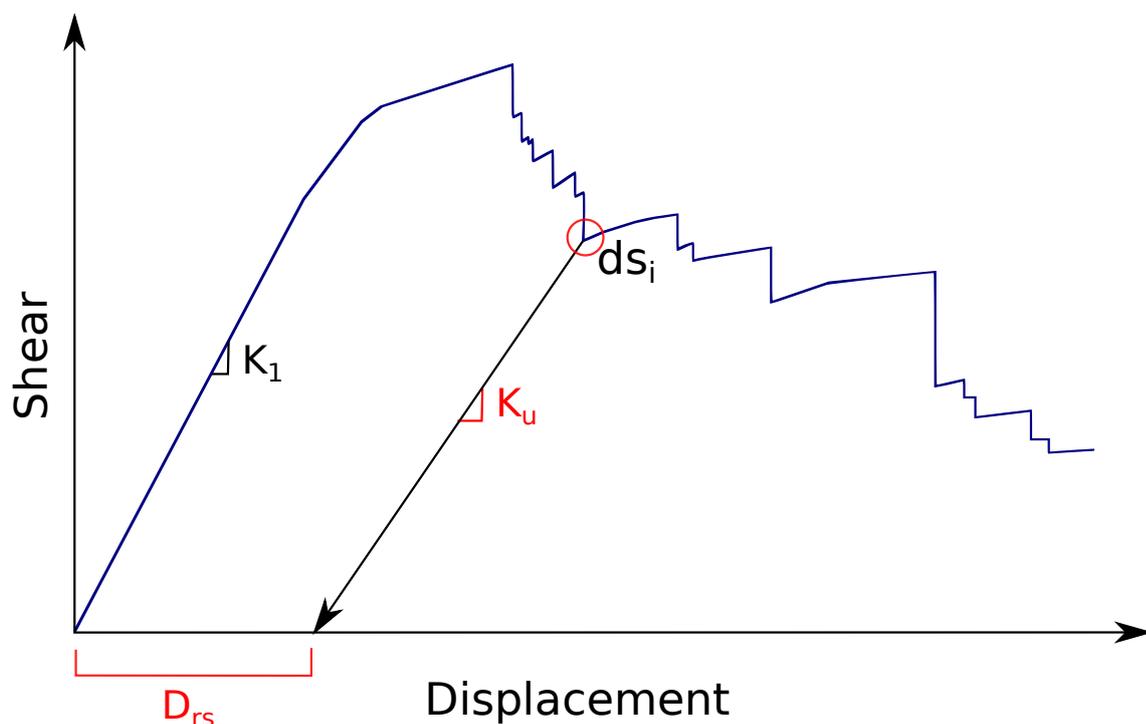


Figure 4.3.: Non-linear static pushover curve for a building, and linear unloading after experiencing a displacement of ds_i with unloading stiffness $K_u < K_1$, and residual static displacement D_{rs} (not to scale; modified from Maffei et al. 2004, p. 16).

oriented towards the rapid reconstruction process in Japan and China, a *pessimistic* model aligned with the reconstruction after the Bam earthquake in Iran (2003; $\ln(\mu) = 1.3$, $\ln(\sigma) = 0.5$ [yrs]), and a model that tends towards a conservative side between these options ($\ln(\mu) = 1$, $\ln(\sigma) = 0.5$ [yrs]).

As time-delayed repairs result in a potentially damaged building stock for subsequent events, the intact building fragility models as described in section 3.2 are no longer sufficient, the exceedance probabilities needing to be defined starting from any damage state. Ideally, the response is estimated using a detailed dynamic analysis, i.e., building an accurate structural model and recording its response during various sets of input ground motion. These probabilities can be determined following, e.g., the approach proposed by Bazzurro et al. (2004). In case a detailed dynamic analysis of a building is not possible, Bazzurro et al. (2004) suggest determining the response using the SPO2IDA (StaticPushOver2IncrementalDynamicAnalysis; Vamvatsikos and Cornell 2006) tool, which estimates the typical distribution of an incremental dynamic analysis (IDA; Vamvatsikos and Cornell 2002) of a single-degree of freedom oscillator (SDOF; with the fundamental period of the building) and considers a deterioration of the building through a provided static pushover curve. IDA (Vamvatsikos and Cornell 2002) means that a building model's response is simulated by exposing it to a set of ground motions scaled to various amplitude levels. The pushover curve defines the capacity of the building, i.e., its deformation resulting from lateral forces (shear; see Figure 4.3). Bazzurro et al. (2004) suggest applying the SPO2IDA tool using capacity curves for the building in the different

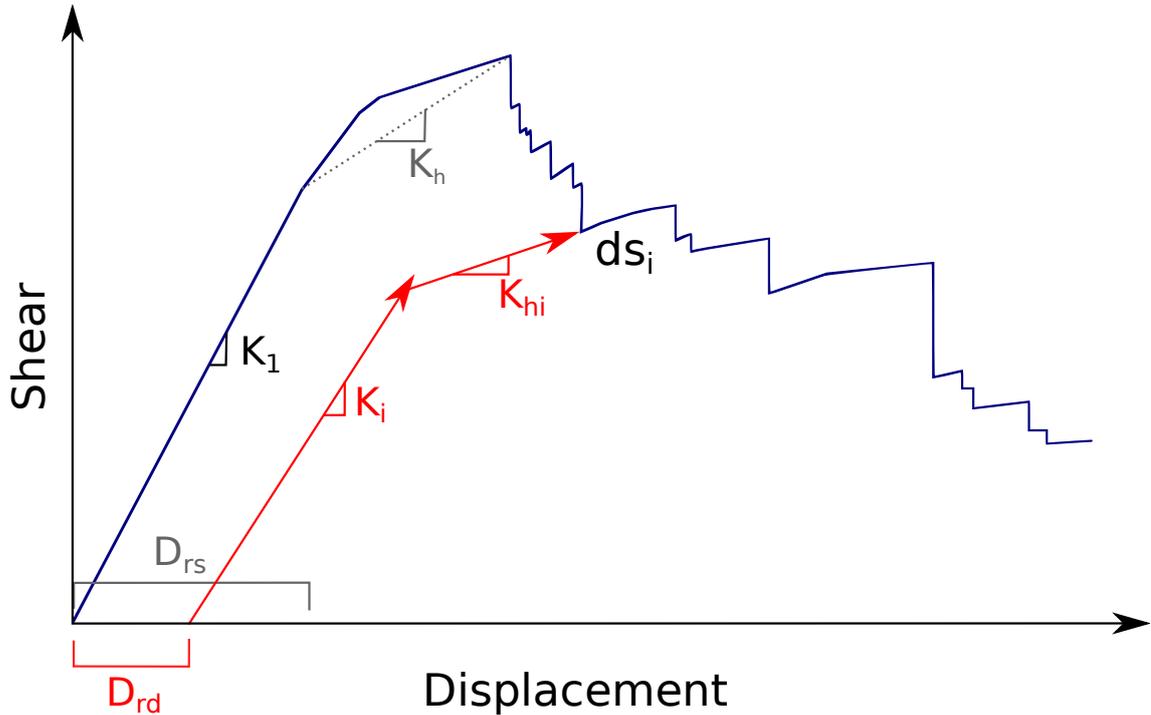


Figure 4.4.: Non-linear static pushover curve for a building, and linear reloading from the residual dynamic displacement $D_{rd} < D_{rs}$ with stiffness $K_i < K_1$ and hardening stiffness $K_{hi} < K_h$, after previously experiencing a displacement of ds_i (not to scale; modified from Maffei et al. 2004, p. 17).

damage states, which can be determined following Maffei et al. (2004). For a building that experienced some displacement d_i , this method assumes that the capacity for displacements $d_j > d_i$ is unchanged, except for a change in the reloading stiffness and a residual displacement. For the damaged behavior, Maffei et al. (2004, p. 26) assume linear unloading of beams and determine for each beam that has exceeded its elasticity an unloading stiffness that depends on the observed damage pattern, i.e., for damaged beams using a cracked moment of inertia ($I_c < I$), $2/3I$ if one beam end fails, and $I/3$ if both failed. From the static residual displacement resulting from this unloading (see Figure 4.3), a dynamic residual displacement and a residual capacity is estimated. The reloading stiffness is then composed of an initial phase corresponding to the unloading stiffness and a hardening stiffness (see Figure 4.4) considering the number of fractured connections of the beam reaching the residual capacity (details in Maffei et al. 2004).

Running the SPO2IDA tool with push-over curves defined in this way for each initial damage state, and fitting lognormal distributions to the quantiles obtained from SPO2IDA at the levels of displacement that define the limit states (see, e.g., equation 3.10), as indicated in Figure 4.5, one can obtain fragility functions starting from each of the damage states.

As the damage process can start, stop and continue at any damage state during multiple events, the individual transition histories for a building are the result of a

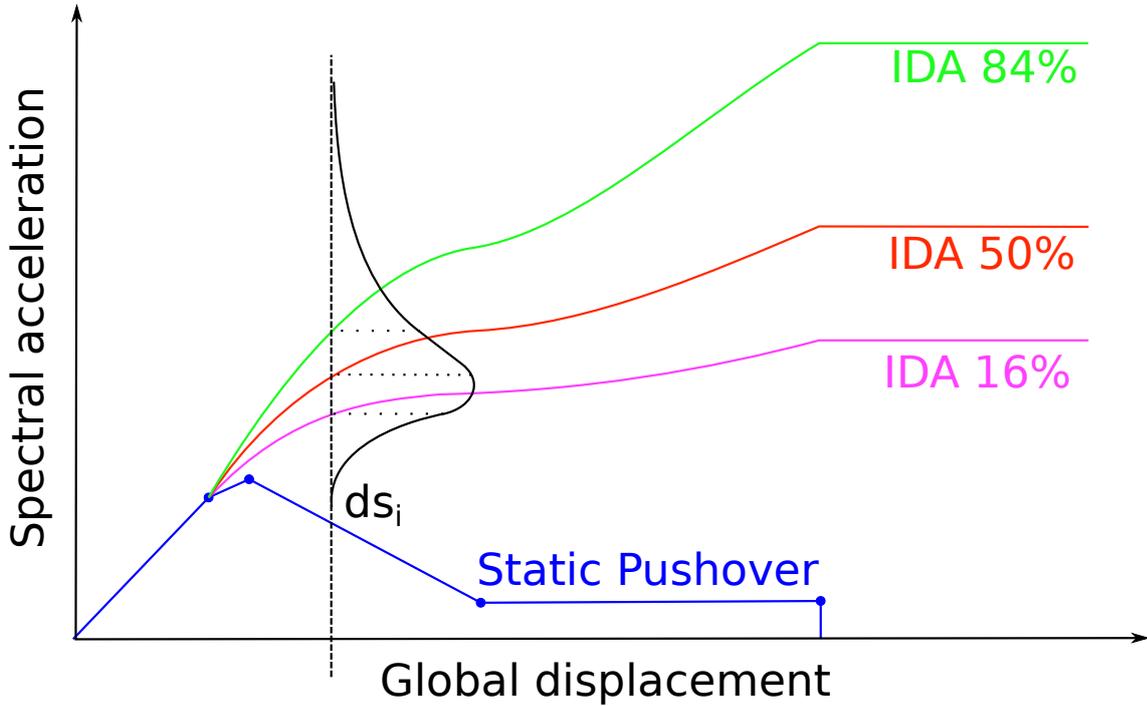


Figure 4.5.: Schematic plot of IDA (Vamvatsikos and Cornell 2002) results (16, 50 and 84% quantiles in green, red, and pink) for a non linear static pushover curve (blue) as returned by SPO2IDA (Vamvatsikos and Cornell 2006) and an indicated lognormal fit (black) describing the fragility for a single damage state ds_i (not to scale; modified from Maffei et al. 2004, p. 22).

stochastic process. Assuming that the probability to go from one damage state to another only depends on the starting damage state and not on the damage history, this process can be described as a time-independent discrete parameter Markov chain (see e.g., Yeo and Cornell 2009b; Iervolino et al. 2016; Trevelopoulos and Guéguen 2016). Formally, this probability to observe damage state x_n is then (Parzen 2015, p. 189):

$$P[X_n = x_n | X_0 = x_0, \dots, X_{n-1} = x_{n-1}] = P[X_n = x_n | X_{n-1} = x_{n-1}] \quad (4.18)$$

given the discrete damage state process X with n realizations, and x_i the transition probabilities.

The transition probabilities x_i are calculated from the the CDF of the corresponding fragility function $F_{i,j}(Y)$ with underlying mean $\mu_{i,j}$ and standard deviation $\sigma_{i,j}$, the transition probability at a given ground motion intensity Y for each damage state ds_i to ds_j can then be defined as a $n \times n$ transition probability matrix, whose elements are the probabilities:

$$P(ds = ds_j | Y, ds_i) = \begin{cases} 1 - F_{i,j+1}(Y) & ; i = j \neq n \\ F_{i,j}(Y) - F_{i,j+1}(Y) & ; j > i \\ 1 & ; i = j = n \end{cases} \quad (4.19)$$

where the last damage state n is an absorbing state, i.e., can't be exceeded.

The approach proposed by Bazzurro et al. (2004) requires a detailed model of the damage patterns to derive the building deterioration. For the purpose of this thesis a simpler approach is followed which while not yielding realistic fragility models, will result in models that are sufficient to show case the application of the proposed simulation framework.

Considering only moment frame buildings without seismic design, a non-linear static push over curve can be determined considering only the columns of the buildings assuming that no seismic damage hierarchy has been considered during the design. For each column, the shear-drift capacity is determined using the assessment described in Eurocode-8 (EC-8) Part III (European Committee for Standardization 2004b). Considering columns as springs in parallel, the story shear capacity is determined as the sum over all column capacities. Considering stories as springs in series and using the Rayleigh method (Clough and Penzien 1975, p. 140), the 2D deformation shape of the building, its base shear vs. roof drift capacity (see Table 5.5), and equivalent single degree of freedom (ESDOF) parameters for the intact structure are determined (see Appendix C.2). Considering the deterioration of the ESDOF through secant stiffness and empirical damping increase (Priestley et al. 2007, p. 86), an IDA (Vamvatsikos and Cornell 2002) is run on the ESDOF system implemented as equivalent column element in the structural response simulation framework OpenSees (McKenna 1997) for each damage state using EC-8 conform ground motions, i.e., a set of at least 7 records, chosen depending on the seismic hazard, the period of the (deteriorated) system (Iervolino et al. 2011) and the EC-8 soil-types on which the building is built (see details in Appendix C.3).

If only a linear ESDOF system is used only linear approximations of the transitions from ds_i to ds_{i+1} are simulated (see Figure 4.6) yielding probabilities $P(ds \geq ds_{i+1} | ds_i, I)$ which are underestimating the fragility. Neglecting parameter changes during the motion, all other transition probabilities can be approximated as:

$$F_{i,j}(I) = P(ds \geq ds_j | ds_i, I) = \prod_i^{j-1} P(ds \geq ds_{i+1} | ds_i, I); \forall j > i \quad (4.20)$$

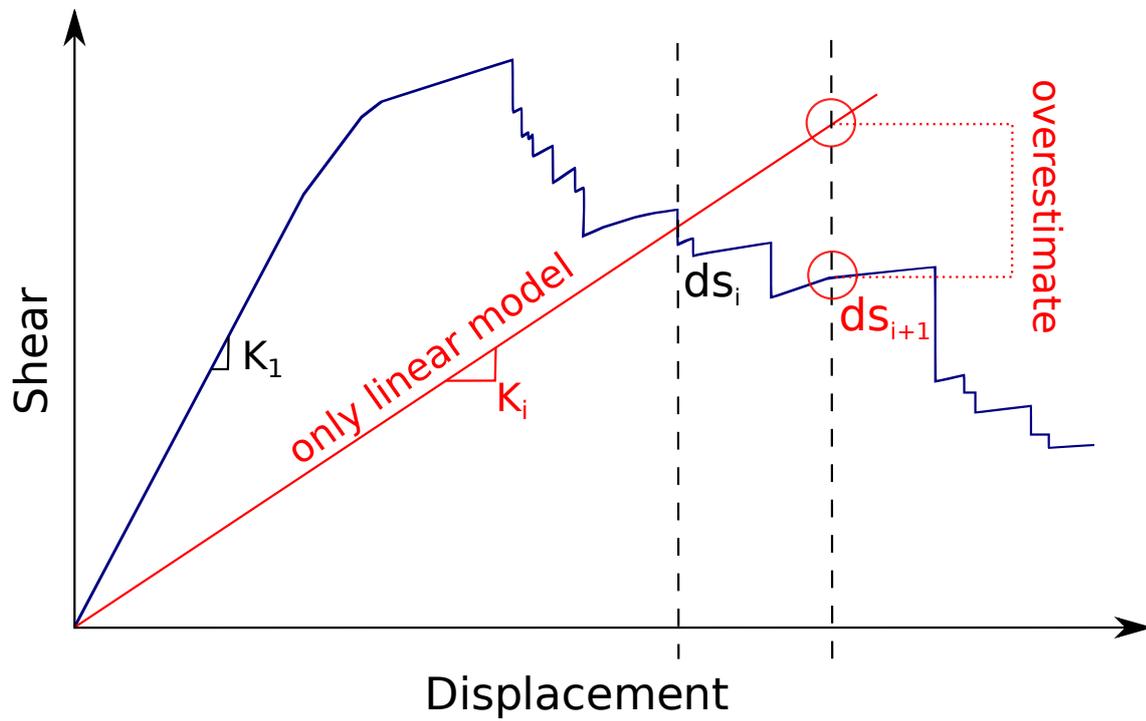


Figure 4.6.: Linear approximation of response yielding a strong overestimation of the capacity and thus an underestimation of the fragility. The resulting fragilities are not realistic, but yield state dependent models suitable for demonstration purposes.

4.3. Implementation of the simulation framework

The framework presented in the previous sections is implemented as a computer program. An overview on its implementation and work flow is shown in Figure 4.7.

A set of simulations is configured through a single configuration file. In the configuration file several choices regarding the individual components and model decisions are specified. These include the number of simulations, the initial random seeds (to produce pseudo random replicable results), the simulation period, the parameters for the ETAS model, the ground motion prediction equation (as available in OpenQuake), the repair function (time dependent or fixed), and switches to consider only Poissonian events, generate full sequences, use only mainshocks of the sequences, and state-dependent or independent fragility. In addition, there are options to reuse previous hazard calculations, or expand upon a previous set of simulations. The work flow consist of three modules (light blue rectangles) dealing with 1) seismicity, 2) ground motion and 3) damage and repair.

The required input (beige rhombuses) are 1) a seismic catalog for the region, alongside its declustered version, 2) an exposure model in the form of locations and building types, and 3) cumulative fragility functions.

For the specified number of simulations, independent realizations of the seismicity within a region for a given observation period are modeled using the proposed approach (first module). Given the regional seismic catalog and parameters b , c , p (equations 4.4-4.10), the framework estimates full seismic sequences resulting from simulated Poissonian seismicity and additional triggering events as defined by a historical input catalog. Since the spatial aftershock placement parameters (γ , l_0 , q , and σ) values (equations 4.4-4.10) are particularly vulnerable to incomplete sequences in catalogs (Hainzl et al. 2014), parameters estimated for densely monitored Southern California (Moradpour et al. 2014) are implemented for the purpose of this study. If a reliable catalog is available, this should be replaced by locally derived parameters. Depth is sampled for a depth distribution defined on a set of points for 2.5, 10, 22.5, and 40 *km*. The rupture mechanism for each event is assigned as defined by the nearest point of a set of points that define a rupture mechanism (e.g., observed from the CMT catalog). For each of the earthquake ruptures, ground motions at locations as defined for a set of buildings are then estimated (second module) using OpenQuake (Pagani et al. 2014). Spatial correlation of ground motion (Jayaram and Baker 2009) is considered as well as $v_{s;30}$ (currently no basin effects) as provided for each of the assets. The history of ground motion is then evaluated as a Markov chain for each individual building employing probabilistic damage and repair functions (third module). Buildings are defined at a point and described by a building type which is related to a (state-dependent) fragility function. Each building can be assigned an (initial) current damage state, which will be updated during the simulation. Alternative fragility models can be provided for each building type and in each independent simulation a single random sampled fragility function is assigned per building type. If a building is damaged, a (potentially-time dependent)

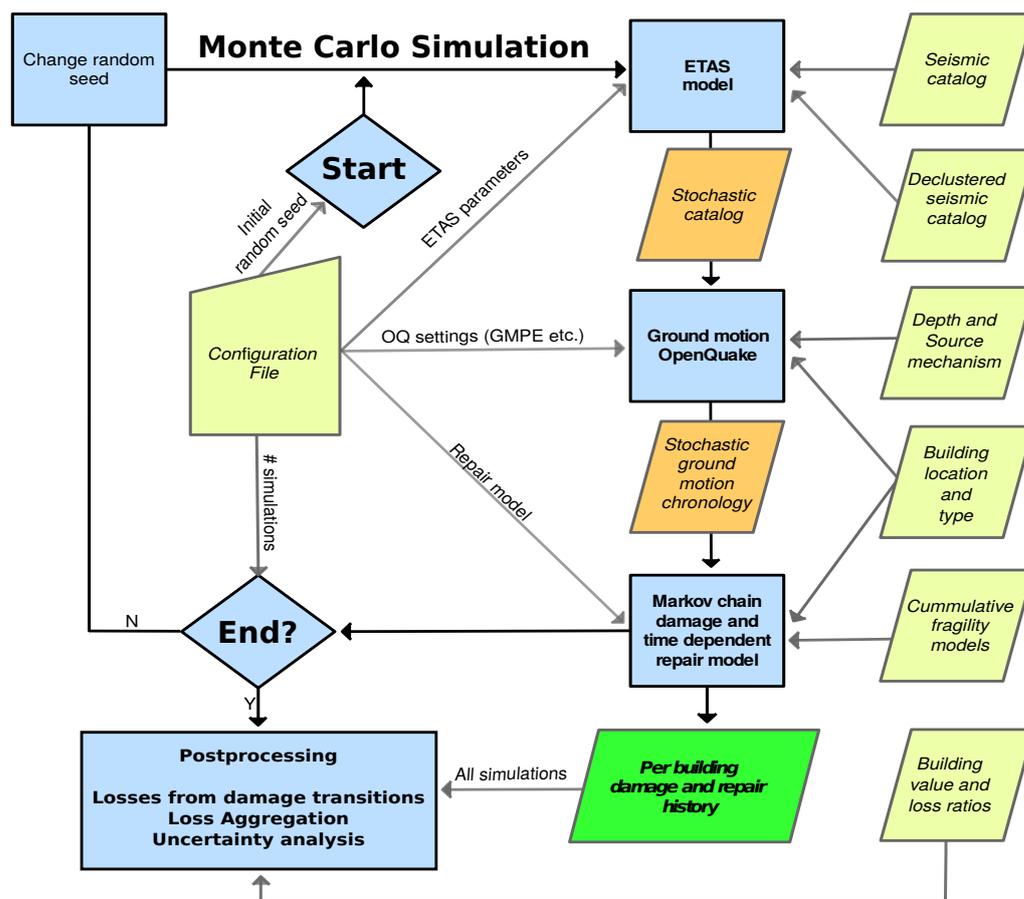


Figure 4.7.: Monte-Carlo simulation framework to model time and state dependent seismic risk over urban scales. The simulation parameters are defined in a configuration file. The required input are a seismic catalog, a spatial model of depths and source mechanisms, location and type of assets, and associated state-dependent fragility curves (see Appendix D for examples). The simulation starts with generating seismicity in the ETAS module and continues anti-clockwise along the work flow until all simulations as defined in the configuration file are done. Results are independent simulations of the damage and repair history for each building.

probabilistic model is evaluated at the time of each considered earthquake event, prior to the damage evaluation, repairing the structure to the intact state in case.

Each simulation yields a history of damage and repairs for all buildings (light green). While the total number of simulations defined in the configuration file is not completed, the framework repeats the simulation, modifying the random seeds systematically. In Appendix E the computational costs of the implementation are briefly evaluated. Except for the alternative fragility models, the implementation currently does not consider any epistemic uncertainty in the components and thus only includes aleatory uncertainty.

Once completed, a post-processing routine allows the estimation of hazard and risk from the simulation outputs. In order to do so, the routine requires the definition

of reconstruction costs and loss ratios for each of the buildings, which are used to translate damages to losses. Employing the boot strap approach (Efron 1979) the routine repeatedly samples the simulations with a user defined sample size and number of repetitions. Calculating for each of the sample set of simulations quantiles of the maximum ground motions and losses, yields hazard and loss exceedance probabilities and their uncertainty as considered within the model.

5. Application to the city of Nablus

The proposed framework was applied to Nablus, which is a major city in the West Bank with an estimated population of 153,061 (Palestinian Central Bureau of Statistics PCBS, 2016). It is situated in a mountainous region along a valley with quite steep slopes (altitudes ranging from 580 m to 940 m; El-Kelani and Jardaneh 2011). The West Bank tectonically belongs to the Dead Sea region, one of the seismically most active regions in the Middle East. The largest tectonic feature is the Dead Sea fault system, expanding from Southern Turkey and the East Anatolian fault zone along the Jordan Valley, the Dead Sea, the Wadi Arava, and the Gulf of Aqaba, until the southern tip of the Sinai peninsula (see Figure 5.1). It is the primary seismogenic structure of the plate boundary between the Sinai sub-plate and the Arabian plate (Salamon et al. 1996). The fault zone is dominated by transform faulting with substantial normal faulting. Several smaller faults expand west and eastwards from the north-south oriented main faults. Fortunately, due to slip rates of only a few millimeters per year between the two plates (4-6 mm/y; Masson et al. 2015), strong earthquakes (i.e., $M_w > 7$) are relatively rare. Nevertheless, based on these slip rates, such an event can be expected approximately once every 250 years. The last earthquake that caused serious damage within the region around Nablus was the 1927 Jericho event. With a moderate magnitude of M_L 6.2 (local magnitude, Avni et al. 2002) it cost about 500 lives in the region and caused widespread damage. While today's elderly are the only witnesses to this event, a recent earthquake in 1995 in the Gulf of Aqaba with a moment magnitude M_w 7.2 reminds of the ever present threat. Luckily, the latter caused limited damage due to its remote location (Hofstetter 2003).

When compared to 1927, the region's population has increased by a factor of more than 20. It is therefore paramount to understand the impact of future earthquakes with magnitudes potentially even larger than the one in 1995, or occurring closer to densely populated areas. Nablus is in close proximity to the Dead Sea Fault Zone which lies less than 30 km to the east of the city. Consequently, it is strongly threatened by seismic events and has a large potential risk. This is even more relevant since an official building code has been introduced only recently (2016) and while most structures are engineered, they are not necessarily designed for seismic loads. As a result, already moderately sized events and aftershocks on smaller faults close to the city might cause severe damage. The city's building stock is relatively homogeneous as it is mainly composed of reinforced concrete frame buildings with infilled walls and few braced frame buildings (Grigoratos et al. 2016, Section 5.2.2). This requires fewer building models to be developed to build a risk model over urban scales. From this perspective, Nablus is a well-suited example to test the proposed methodologies. Please note that several (fairly strong) simplifications are

made during the application of the framework here and risk estimates presented later on are only indicative and not intended to represent reliable estimates but only intended to show case the frame work. Regardless of this, each component will be briefly evaluated for its validity as far as possible. However, since no recent damaging events have been observed, empirical data for validation are missing.

5.1. Seismic hazard for Nablus

5.1.1. Seismicity model

To derive the seismicity model, a catalog of historical earthquakes was compiled for the Dead Sea region, integrating data from different sources (see Appendix A.1 for details). Figure 5.1 shows the resulting catalog which spans from the year 23 until 2014, was harmonized to moment magnitude and includes 14'510 events above M_w 3. The simulation framework employs a simple gridded source model as described in the previous session. Currently, the only input to the probabilistic source model generation is thus a declustered version of the catalog. It is declustered using the time-window definition of Grünthal as reported by van Stiphout et al. (2012) and the Gardner and Knopoff (1974) moving window algorithm. Catalog completeness is estimated as shown in Table 5.1. using the visual approach as described in Burkhard and Grünthal (2009) with 0.5 M_w magnitude bins.

Excluding events before 1990 or below M_w 4 (m_{min}), a single regional maximum likelihood b-value is estimated using equation 3.6. Considering a rounding error of 0.1 in the catalog magnitudes, running a Monte-Carlo simulation to generate a stochastic set of catalogs and repeatedly estimating b yields $b = 1.12 + / - 0.04$ after 10'000 iterations ($a = 6.14$; $M_c = 4.2$). From the full catalog, a maximum magnitude of M_w 8 is estimated using the Bayesian estimator by Kijko (2004).

For the earthquake source model generation, a grid with a 0.1 degrees discretization and a Gaussian kernel with a width of approximately the average spatial uncertainty for the instrumental period (see Table A.2) $\sigma = 20km$ was chosen (cut at $3\sigma = 60$ km) for the Frankel (1995) approach. Figure 5.2 shows the resulting spatial probability for Poissonian events.

The hypocentral depth distribution is estimated from catalog events since 1900. As expected, for a mainly transform fault regime with a crustal thickness of around 32-37 km (Hofstetter et al. 2000), almost 80% of these events occurred at a depth

Table 5.1.: Year of completeness estimates for 0.5 Mw bins using Burkhard and Grünthal (2009)

Year	1987	1963	1945	1895	1870	1000	330
Mw	4.25	4.75	5.25	5.75	6.25	6.75	7.25

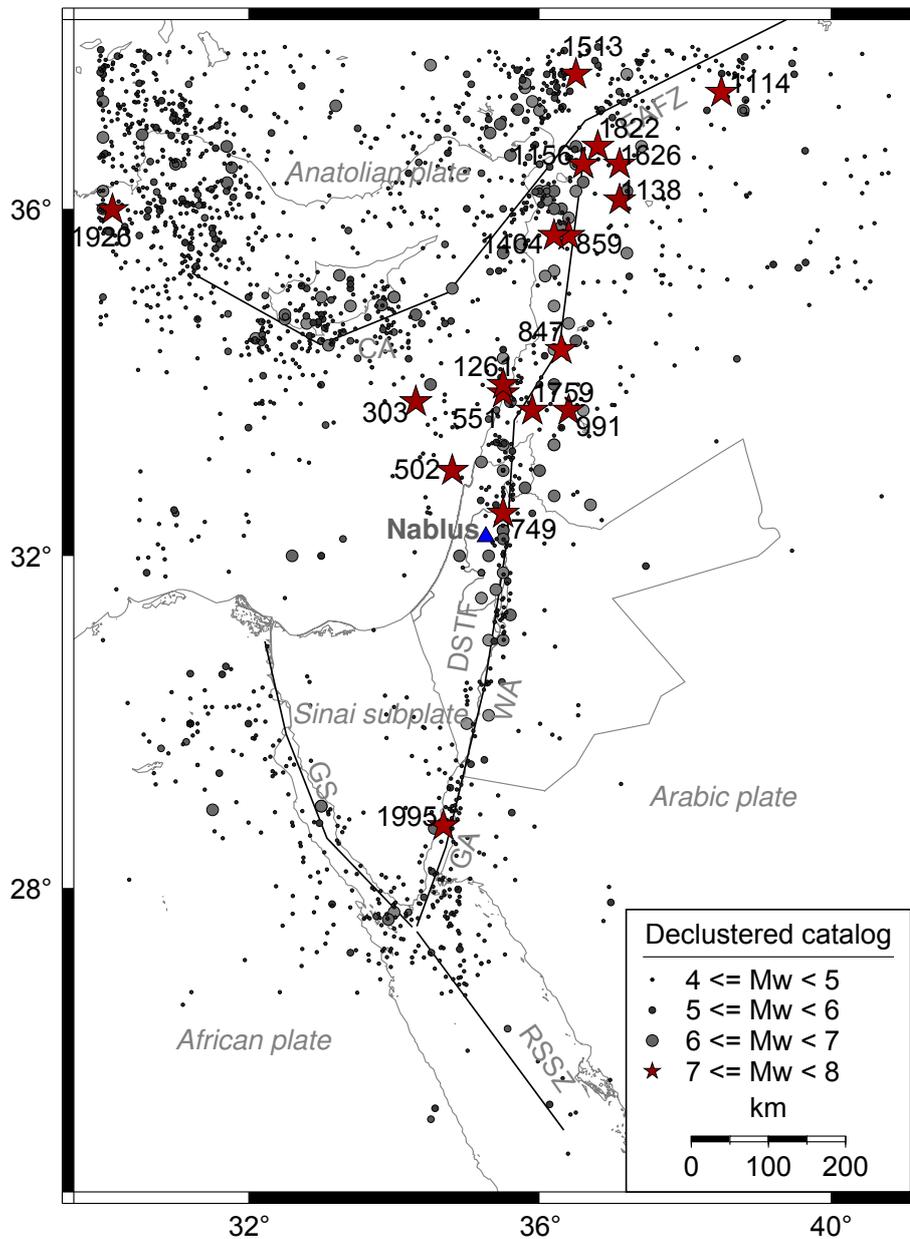


Figure 5.1.: Location of Nablus in the Dead Sea region (triangle), and events ($M_w \geq 4$) contained in the declustered catalog window and harmonized to M_w (circles, stars indicate strong events annotated with their year of occurrence). The main features of the region's tectonics (black solid lines, modified from Salamon et al., 1996) are also shown: RSSZ, Red Sea spreading zone; GS, Gulf of Suez; GA, Gulf of Aqaba; WA, Wadi Araba; DSTF, Dead Sea transform fault; CA, Cyprian arc; EAFZ, East Anatolian fault zone (Figure modified from Haas et al. 2016b)

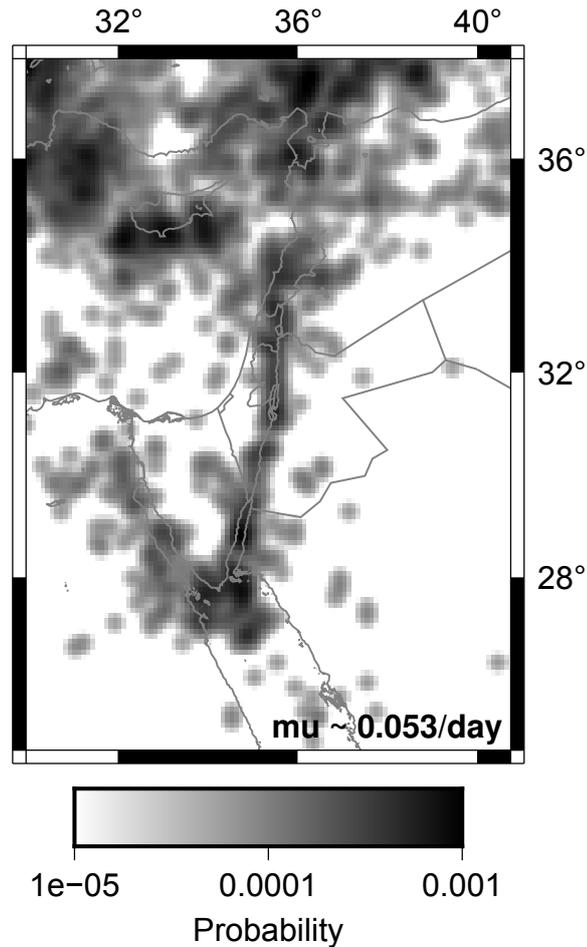


Figure 5.2.: Spatial probability model for the occurrence of Poissonian seismicity obtained after normalizing the gridded a -values resulting from the Frankel (1995) approach.

shallower than 50 km. Most, i.e., three quarters of the deeper observed events are located close to the subduction regime of the Cyprian arc on the western side of the East Anatolian Fault Zone. Thus, for reasons of simplicity and assuming that the region of interest is mostly endangered by shallow earthquakes, the model is restricted to a depth of 50 km. The depth distribution is then estimated on a 4 by 4 degrees grid distinguishing events until 5, 15, 30, and 50 km depth, where bin centers are used as the model depths (Table 5.2).

The source mechanism for each event is assigned based on the few historical data that are available for the region. The Global Centroid Moment Tensor database (CMT; Dziewonski et al. 1981; Ekström et al. 2012) lists 54 events for the considered region. A simplified approach was followed, where, for each of the double couple solutions, the most likely fault plane is determined as the one whose strike direction deviates the least from the strike direction of the closest segment of a major fault, as described by Garfunkel et al. (1981). Figure 5.3 shows the CMT events and the major faults. The chosen planes and event locations are then used as a simple spatial model for the source mechanism.

For the simulation framework, parameters for the ETAS model have to be pro-

Table 5.2.: Depth distribution determined from the 100 events closest since 1900 to a set of grid points (longitude, latitude).

Longitude ¹	Latitude ¹	0–5 (km)	5–15 (km)	15–30 (km)	30–50 (km)
31.95	26.55	0.18	0.54	0.19	0.09
31.95	30.55	0.10	0.52	0.27	0.11
31.95	34.55	0.04	0.35	0.35	0.26
31.95	38.55	0.11	0.39	0.20	0.30
35.95	26.55	0.23	0.65	0.10	0.02
35.95	30.55	0.18	0.67	0.11	0.04
35.95	34.55	0.23	0.53	0.12	0.12
35.95	38.55	0.15	0.51	0.14	0.20
39.95	26.55	0.22	0.65	0.09	0.04
39.95	30.55	0.19	0.59	0.14	0.08
39.95	34.55	0.16	0.58	0.15	0.11
39.95	38.55	0.14	0.59	0.11	0.16

¹ Cell centers of a 4° by 4° grid over the catalog region.

Table 5.3.: ETAS parameters. Values are taken from Moradpour et al. (2014) except if marked with *, which are derived in the scope of this study.

α^*	b^*	c^*	d	γ	K^*	l_0 [km]	μ^*	p	q	σ
1.12	1.12	0.003	1.2	0.6	0.0025	0.01	0.0053	1.1	0.35	0.44

vided. As described previously values from Moradpour et al. (2014) are used for the aftershock placement (γ , l_0 , q , and σ), and in addition also the exponential decay with time $p = 1.1$. It is assumed that this choice is acceptable here, since Southern California is to some extent tectonically similar to the region around the Dead Sea. Both are dominated by mainly transform faulting, but California also shows reverse faulting. The average slip-rate is smaller for the Dead Sea (around 4-7 mm Masson et al. 2015) with respect to Southern California (5-37mm Johnson 2013), and this is captured by deriving the local Gutenberg-Richter parameters (daily rate μ , b) and the productivity K . The parameter for the time delay in the onset of sequences according to the modified Omori law is set to $c = 0.003days$ ($c(m = m_{min}) = 10^{m-6.5}days; m_{min} = 4$). The angle from the mainshock at which the aftershock occurs is chosen randomly, whereas the radial distance is chosen according to equation 4.10, where the distance density was discretized at 10m. All considered ETAS parameters are presented in Table 5.3.

Figure 5.4 shows a single simulated catalog with the triggering background events and the triggered events. Please note that only events with $M_w \geq 4$ are considered.

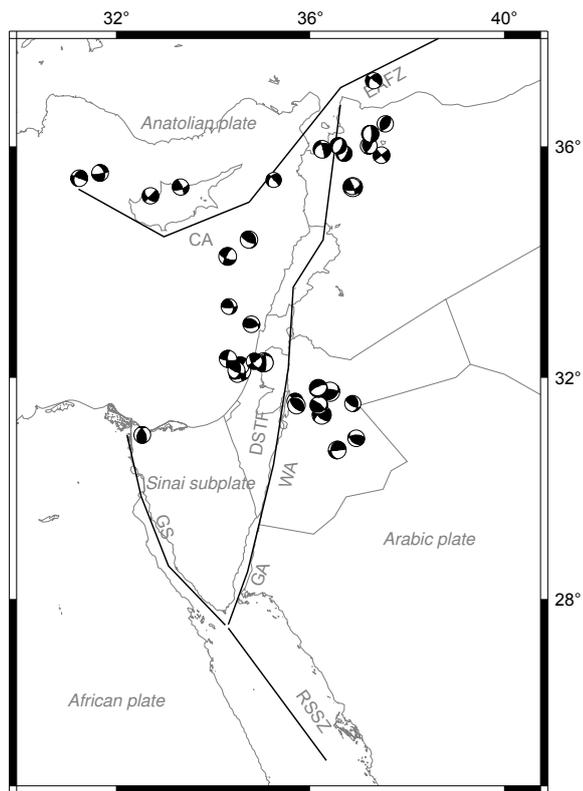


Figure 5.3.: Events found in the Global Centroid Moment Tensor database (CMT; Dziewonski et al. 1981; Ekström et al. 2012) for the region of interest and a simplified description of the tectonics (Garfunkel et al. 1981)

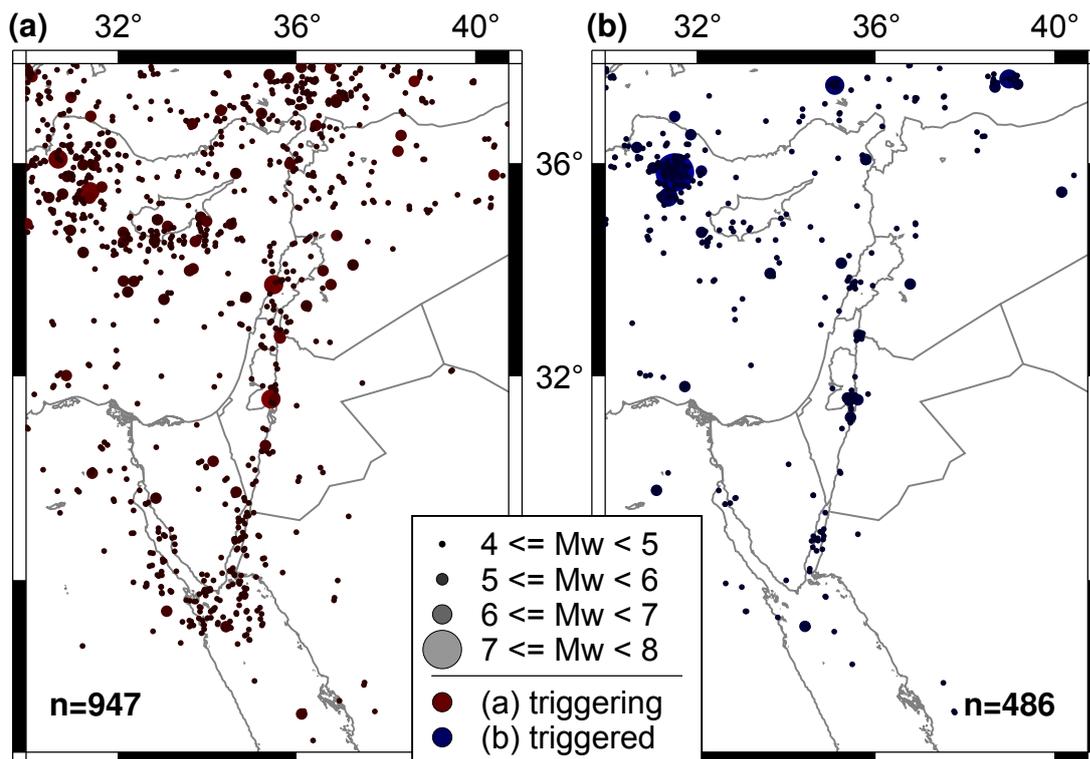


Figure 5.4.: Example for a typical simulated catalog over 50 years; (a) shows the background events obtained from the Poissonian model and (b) the events triggered by the events in (a) according to the inhomogeneous Poisson process as described by the ETAS model. Note that triggered events may include larger events (main shocks) and only events with $M_w > m_{min}$ are plotted here.

5.1.2. Ground motion models

As strong motion data in the region are sparse (e.g., RESORCE database: 26 records for 8 earthquakes in Egypt, Israel, Lebanon and Syria; Akkar et al. 2014), this hinders robust statistical tests for the selection of GMPEs. Thus, only the GMPE developed by Campbell and Bozorgnia (2014) is used. This model supersedes the Campbell and Bozorgnia (2008) model which had previously been found suitable for the region (Klar et al. 2011; Levi et al. 2015). Furthermore, only PGA will be evaluated, although other parameters like spectral acceleration show a better correlation with damage, especially for larger structures. Nevertheless, PGA is chosen as it requires the calculation of only one value, rather than values for a range of periods for different buildings. However, the implementation of the framework could be extended to other measures in future.

5.1.3. Site-effects for Nablus

For a risk assessment over urban scales site effects are a crucial component. Since the valley of Nablus is covered by Tertiary and Quaternary sediments of gravel deposits and clay sediments (El-Kelani and Jardaneh 2011), one can expect potentially high levels of ground motion amplification. Unfortunately, the geotechnical analyses for Nablus available in literature are sparse (Jardaneh 2004; Jardaneh 2007; Faidi 2007; Al-Dabbeek and El-Kelani 2008; El-Kelani and Jardaneh 2011). Most of these rely on single station noise measurements and yield only the fundamental frequency of the soil applying the Nakamura (1989) horizontal to vertical (H/V) spectral ratio method. Only Jardaneh (2004) provides a single idealized shear wave profile for Nablus. They reported soil types as weathered clay or marl, marly limestone and gravel, and even massive limestone. The top layers range between 0.5 and 1.5 *m* in the western part of Nablus, reaching up to 16 *m* thickness in the center and 5 to 7 *m* in the eastern parts of the city (El-Kelani and Jardaneh 2011). Jardaneh (2004) proposed a single profile with velocities varying between 125 and 159 *m/s* for Nablus. A study questioning the general applicability of v_{s30} to estimate amplification in the region (Zaslavsky et al. 2012) found for similar materials in the region much higher shear wave velocities of 300-600 *m/s* for clay, 650-800 *m/s* for marl and more than 1100 *m/s* for limestones. This raises doubts that the reported profile is representative for whole Nablus.

Lacking spatially distributed and reliable estimates, v_{s30} values inferred from topography are employed (USGS; Wald and Allen 2007; Allen and Wald 2009). This is a critical point, and usually strongly discouraged over urban scales and should be avoided in actual applications, as many studies have shown strong discrepancies between the inferred and measured v_{s30} (see e.g., Mucciarelli and Gallipoli 2006). Figure 5.5 shows v_{s30} values inferred this way for Nablus. For simplicity, only one random realization of the ground motion field for each event is estimated, with (isotropic) spatial correlation of the intra-event residuals according to the Jayaram and Baker (2009) model and clustering, i.e., spatial correlation of v_{s30} , are considered.

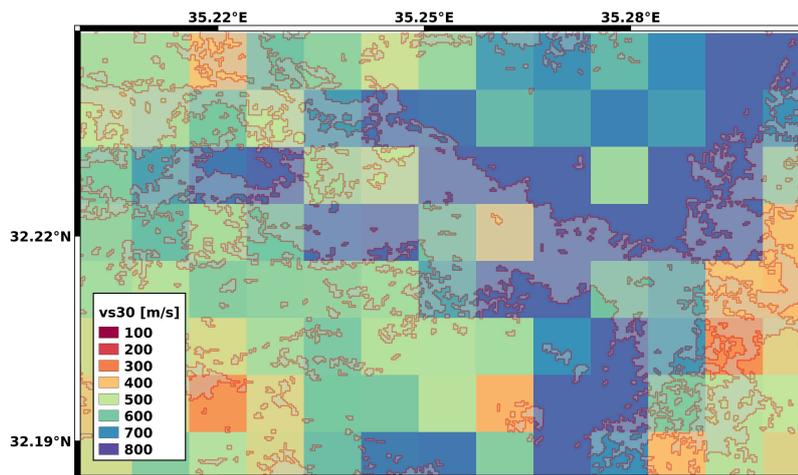


Figure 5.5.: Built-up area of Nablus (gray) and v_{s30} values inferred from topography with a resolution of 30 arc seconds ($\approx 800m$; Wald and Allen 2007; Allen and Wald 2009), used as a proxy for the site-effect.

5.1.4. Resulting hazard estimates

The simulation framework does not provide probabilistic seismic hazard curves directly, but they can be estimated as described in Beauval et al. (2006). This procedure, implemented also in the postprocessing routine, calculates the percentiles of the distribution of maximum PGA observed at a site for each 50 years period simulation, which yields the 50 years hazard curve for the respective site. Figure 5.6 shows the median hazard curves (solid lines) resulting for an example site in the center of Nablus (35.268 E, 32.220 N, $v_{s30} = 760m/s$) using 1'000 bootstraps with a 10% random sampling of the ETAS simulations. Also indicated are 5/95% (dotted), and 25/75% (dashed) quantiles, calculated from the bootstraps. On the right side, the hazard for the same site is shown using simulations that exclude fore- and aftershocks and only consider the events generated by a Poisson process. Please note that these are not equivalent to the mainshocks of the simulation, but the first event in each sequence. However, as both are a result of the same Poisson process, they should converge with an increasing number of simulations. Still, due to the pseudo random simulations here, the comparison might be biased (Please refer to a brief analysis of this effect on the losses in Chapter 6). While the two curves in Figure 5.6 correspond to a site with a relatively high estimated $v_{s30} = 760m/s$, Figure 5.7 shows a site close to the Balata refugee camp in the East of Nablus (35.293 E, 32.206 N) with a lower estimated $v_{s30} = 327m/s$.

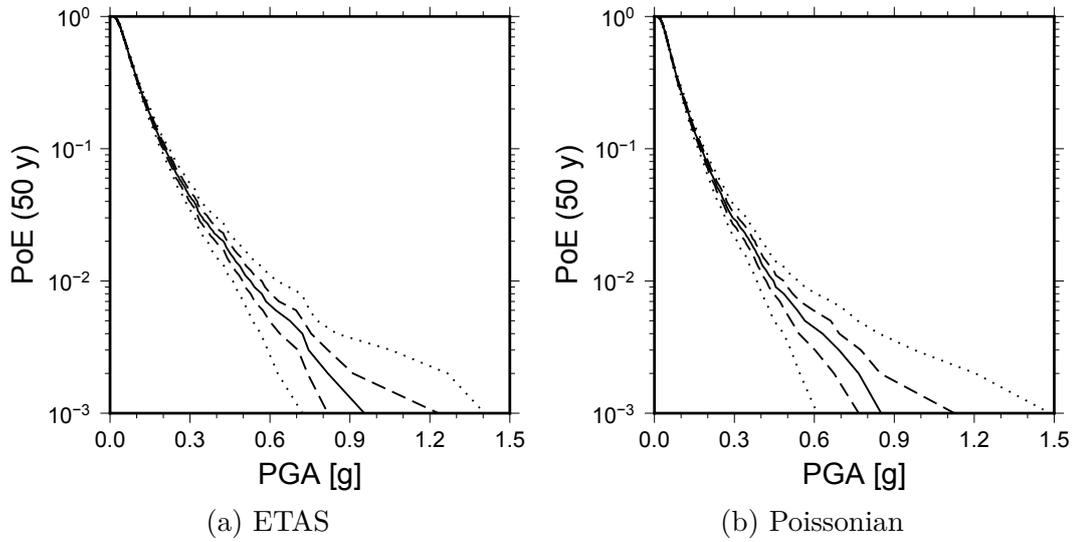


Figure 5.6.: Median hazard curves (solid lines) obtained for a single site in central Nablus (35.268 E, 32.220 N, $v_{s30} = 760m/s$) determined using a 10% bootstrap sampling approach alongside the 5 and 95% quantiles (dotted) and 25 and 75% quantiles (dashed). The left side shows curves resulting from using the ETAS model and the right side using only a Poissonian model.

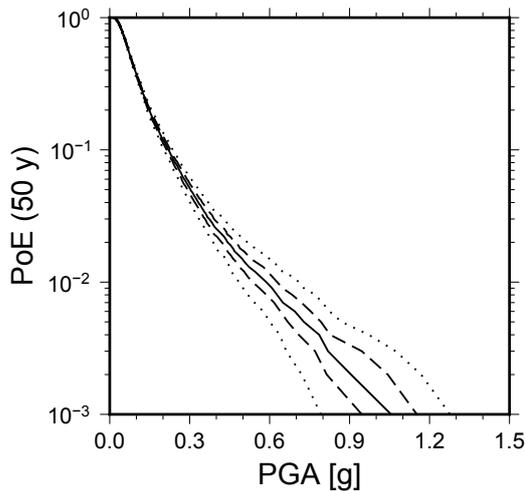


Figure 5.7.: Median hazard curve (solid line) obtained for a single site in eastern Nablus (35.293 E, 32.206 N, $v_{s30} = 327m/s$) determined using a 10% bootstrap sampling approach alongside the 5, 25, 75 and 95% quantiles (dotted, dashed, dashed, dotted, respectively).

5.2. Exposure of Nablus

In order to collect exposure data for the city of Nablus, the general census of the Palestinian territories in 2007 (Buildings and Housing Units Census 2007) can be used. It contains information about population and housing by locality. According to this census, the total population of Nablus in 2007 was 126'131 and the number of buildings was 10'415, with 32'018 housing units and an average household size of 5 people, considering 25'104 households. Recent projections by the Palestinian Central Bureau of Statistics (PCBS) estimates Nablus to have a population of 153'061 in 2016, an increase of about 21% since 2007. Using this number, one can estimate an approximate expected number n of buildings in Nablus for $t_1 = 2016$ using the simple relation:

$$n^*(t_1) = \frac{m^*(t_1)}{m(t_0)}n(t_0) \quad (5.1)$$

where $n(t)$ is the number of buildings in the year t , $m(t)$ is the number of inhabitants and variables with asterisk correspond to an estimate rather than observed numbers, and $t_0 = 2007$, which yields $n^*(2016) = 12'639$. Note that this relation only holds if the household size and the number of households per building are constant over the projection period. Thanks to a footprint dataset of 2016 provided by the GIS and IT Department of the Municipality of Ramallah in the West Bank (personal communication S. Aldwaik), the sanity of such an estimate can be checked. Using the same census data sources for Ramallah, one obtains with $m^*(2016) = 35'140$, $m(2007) = 27'460$ and $n(2007) = 3'086$ an expected number $n^*(2016) = 3'949$. Comparing this estimate to the number of footprints in the dataset, $n(2016) = 4'495$, an underestimation of 12% is observed. Considering that Ramallah is in a unique condition being the de facto administrative capital of the Palestinian National Authority since the Oslo Accord in 1993, and thus has a high density of governmental buildings, NGOs, hotels etc., one expects Ramallah to be on the high end. Thus, one can define a uniform distribution for the building numbers in Nablus as $U(0.88n^*(2016), 1.12n^*(2016)) = U(11'122; 14'155)$. Using the same scaling as in equation 5.1 for the number of housing units, a value of 38'854 units is estimated. This value is similar to the value of 36'574 ($\sim 6\%$ greater) obtained in a recent study on Nablus by Grigoratos et al. (2016).

Three additional variables giving useful insights in the building stock were collected in the Palestinian census. These are: 1) material of external walls, 2) construction period and 3) number of floors of the buildings. While the first two are available on district scale, i.e., the Governorate of Nablus, the latter are aggregated only for the whole Palestinian territory. Unfortunately, the census does not provide any cross tables for these variables. According to this information, one should expect structures with walls made from cement blocks (38%), stone and concrete (16.3%), concrete (16%), cleaned stone (12.6%), old stone (12.3%), adobe clay (0.3%), other (2.8%) and unknown (1.8%) materials. Few buildings were built before the Arab-Israeli war in 1948 (12%), a larger portion under Jordanian rule until 1967 (25%), 15 % in the first 10 years after the six-day war (until 1977), 25% between 1978 and 1987, 10 % between 1988 and 1997, and 11% from 1998 until 2007, the remaining 2% are of an unknown age. Only 17 % of the buildings should be higher than 3

storeys, 39% should have 2, and 42% only one storey (1% unknown).

Theoretically, this information could be used as a rough exposure model for Nablus. The associated uncertainties though are very high, first due to the low spatial resolution and second due to ambiguity in the case of building materials and missing conditional frequency information (e.g., cross tables).

5.2.1. Data collection

In order to reduce this uncertainty, a ground-based survey was carried out to collect panoramic imagery. Using the simplified satellite analysis (Appendix B) on a Landsat 8 scene of Nablus from the 10th of July 2015 yields a land-use and land-cover classification. Please note that only ground truth visually derived from high-resolution satellite imagery (Google; DigitalGlobe 2017) was used here and thus the level of detail is limited, but could be improved if better ground truth is available (Wieland et al. 2012). Combined with a built-up/non-built-up classification of Landsat imagery from 1972 and 1992, this yields strata as shown in Figure 5.8.

These results can already be used to constrain models obtained from census data (Haas et al. 2016d; Haas et al. 2016a). In the multi-stage analysis framework followed here, they can provide the necessary strata to drive additional ground-based exposure data sampling. Assuming the results from the remote sensing analysis yielded a stratification of the town building stock, representative samples of the building stock can be drawn. Figure 5.9 shows the route that was planned for Nablus with a maximum length of 70 *km*.

In July 2016, the route was followed using the GFZ mobile mapping system (GFZ-MoMa). Figure 5.9 shows the path followed in comparison to the planned route. Within a few hours, 90 GB of geocoded panoramic images were collected with a 6 Hz sampling along the path and covering about 2'000 buildings in Nablus.

The images acquired by the mobile mapping campaign were analyzed using the remote rapid visual screening technique (RRVS, Wieland 2013). For this purpose, building footprints were collected from OSM and the images resampled at a spatial distance of 10 meters, and stored on a server alongside the footprints and an associated PostgreSQL/PostGIS database which implements the GEM taxonomy (Brzev et al. 2013). The images and the database can be accessed and interacted with via a simple web application (the RRVS web tool). Using a distance buffer of 30 meters around the collected images, approximately 3'150 footprints were deemed screenable. Of these 940 were screened by local engineers through the web platform (see Figure 5.10), where 827 of these could be analyzed for building features. Out of the 113 supposed buildings (OSM footprints) that could not be screened, 5 did not correspond to an actual building, 91 where not visible from the panoramic image, 12 where not clearly visible, and for 5 no reason was given as to why they were not screened.

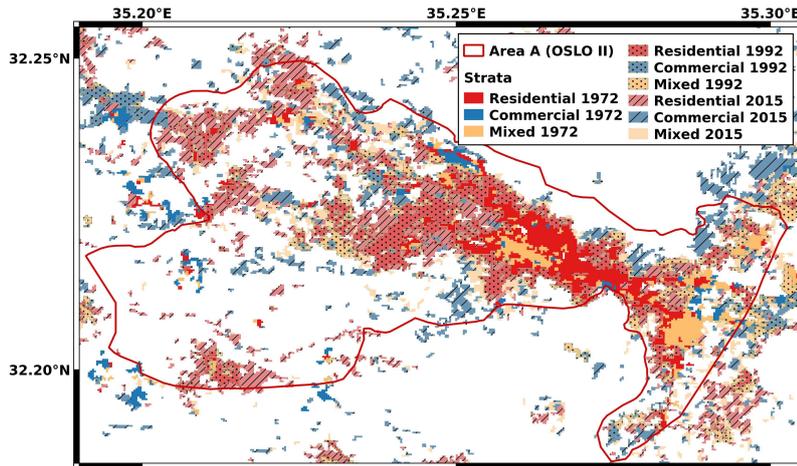


Figure 5.8.: Resulting built-up strata using a pixel-based satellite imaging analysis (Appendix B) distinguishing residential, commercial, and mixed landuse in 2015 and approximate period of urban development (< 1972, 1972-1992, 1992-2015).

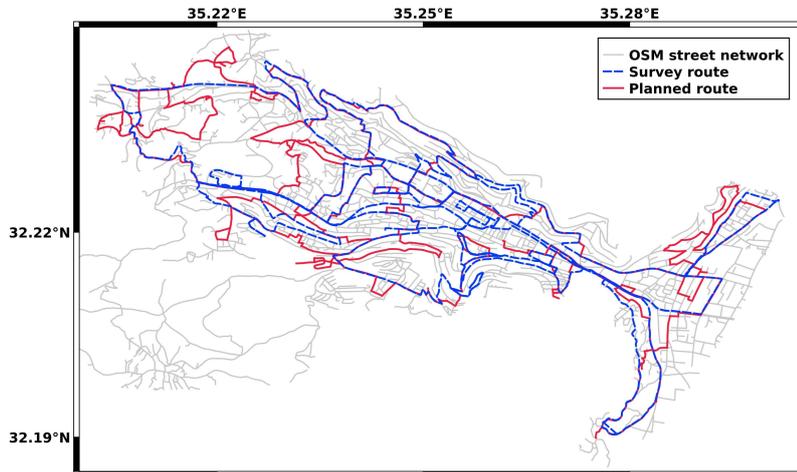


Figure 5.9.: Route resulting from randomized stratified sampling approach, where randomly sampled nodes from the street network fall into the different strata.

Figure 5.11 shows a comparison between wall material as indicated in the Census compared to the exterior wall data interpreted using the RRVS. With about 10%, the share of masonry walls in the RRVS data is much lower than expected from the census. Excluding *Stone and concrete*, which corresponds to 16% in the census, *Adobe Clay*, *Old stone* and *Cleaned stone* yield in total 25% masonry. This large difference can be related to the fact that the census also includes rural areas, but can also result from misinterpretation in either the RRVS or census data.

Figure 5.12 shows the load-bearing material as identified through the RRVS. The clear difference, i.e., between this (i.e., 95% concrete buildings) and the exterior wall materials in 5.11, demonstrates that the frequently used procedure of mapping buildings by the exterior wall materials as indicated in the census (16% *Concrete*, and 16% *Stone and Concrete*) to structural building type is not straight forward

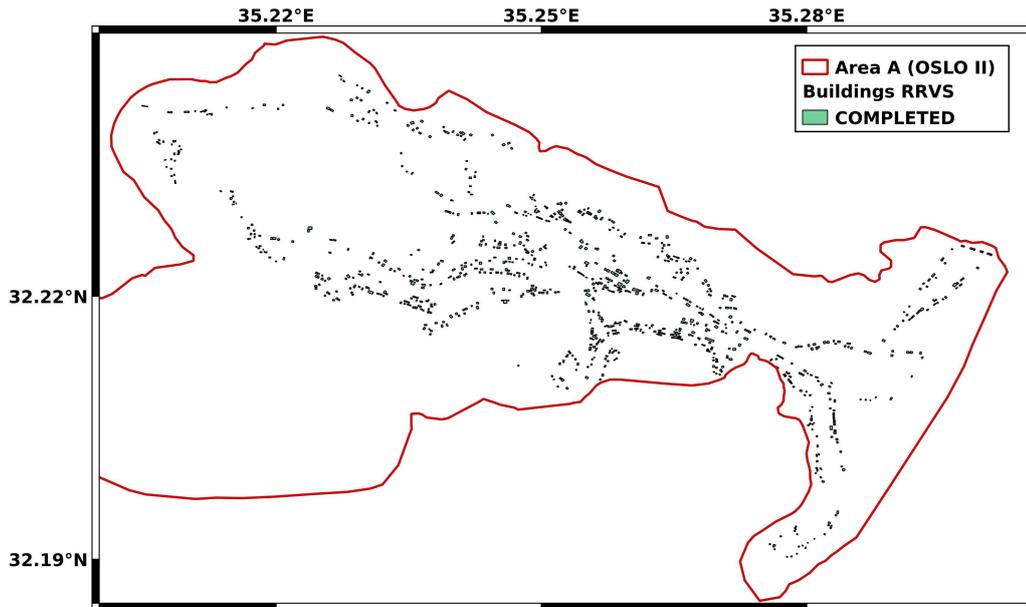


Figure 5.10.: Buildings screened by local engineers via the Remote Rapid Visual screening (RRVS) system using panoramic imagery collected along the path shown in Figure 5.9.

and could result in strongly biased models.

Figure 5.13 shows cumulative density of year of construction as defined in the census for whole Palestine, in comparison to the years assigned to the buildings in Nablus through the RRVS. Although it is not easy to determine building ages visually, the large differences between both distributions might stem from an actual difference for Nablus when compared to the rest of Palestine.

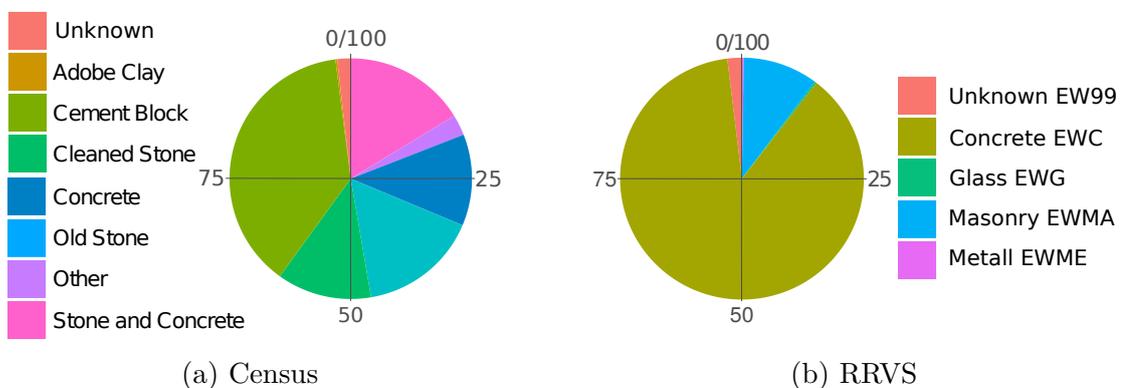


Figure 5.11.: Comparison for exterior wall materials as indicated by the census (on the district level) and the data derived from imagery collected with the GFZ-MoMa and interpreted with the RRVS (city level).

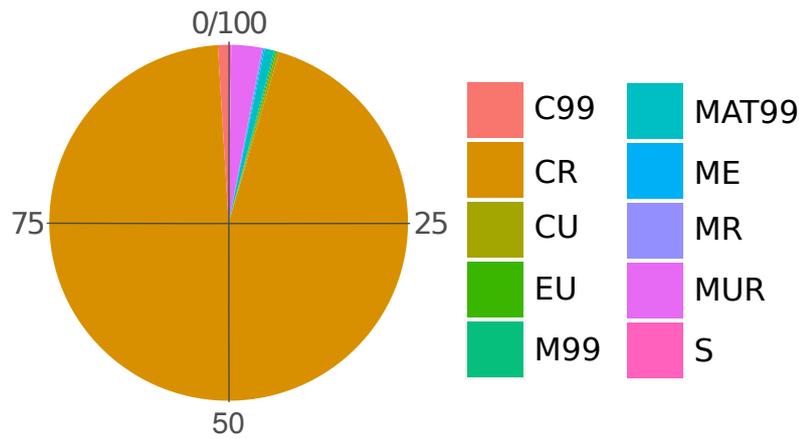


Figure 5.12.: Structural material types in Nablus as observed through the RRVS; Concrete reinforced (CR), unknown reinforcement (C99), unreinforced (CU). Analogous: Masonry (MR, MU, M99); Earthen unreinforced (EU); Steel (S) and unknown (MAT99).

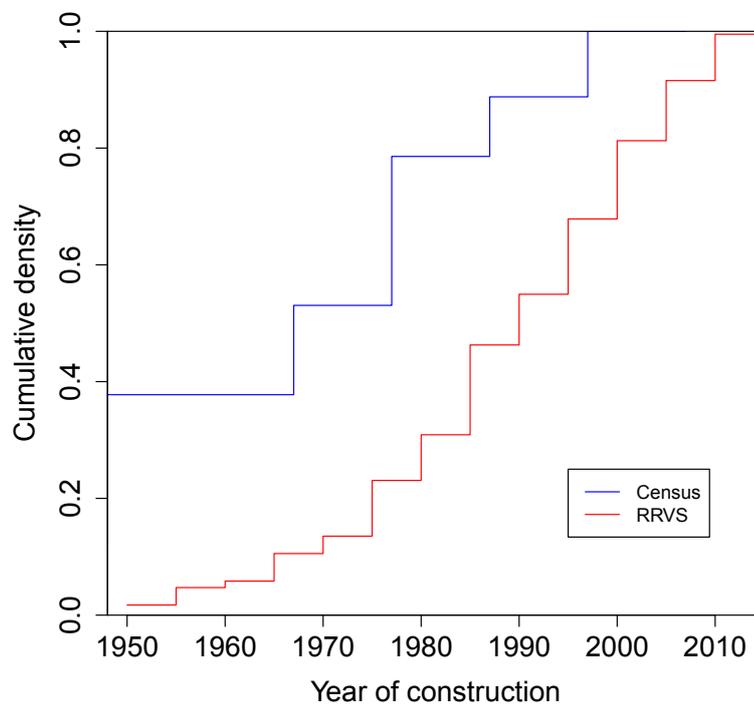


Figure 5.13.: Cumulative density for the year of construction of buildings as indicated in the census (for the whole of Palestine) and as interpreted during the RRVS (Nablus). Even if considering the difficulties to assign building age during visual screening, the differences are quite large.

5.2.2. Exposure model for Nablus

The building data collected within the strata corresponds to about 6.5% of the buildings in Nablus. This data can be either used to a) reduce uncertainties of a census-based model, employing e.g., Bayesian updating (Wieland 2013), or b) being input for a bottom-up model based solely on the collected data. Among the 827 buildings, 714 have a unique attribute combination (counting unknowns as an attribute value). If these are used as individual building types for the fragility analysis, the sample is not sufficient to obtain a reliable distribution for the whole city.

As described in section 3.3.2, there are several ways to group the building data. For demonstration purposes of the risk framework and given the homogeneity of the data, the collected data is not fully exploited and only a simple mapping from observed attributes to types is used. The data collected during the survey suggests that about 96.5% of the buildings are concrete (0.2% unreinforced, 1% unknown reinforcement), 2.8% unreinforced masonry, and 1.7% other materials which were only observed once or twice in the sample (metal: 2, reinforced masonry: 1, steel: 1, unreinforced earthen: 1, masonry with unknown reinforcement: 1). Of the reinforced concrete (RC) buildings, 75.9% are infilled frame and 23.1% braced frame buildings. Other systems were rarely observed (moment frame: 6, hybrid: 1, dual: 1). In agreement with the fact that a seismic code has been introduced only after the collection of the data in 2016, the total number of ductile buildings among the RC buildings is very low (2.9% RC ductile of RC, 0.1% base isolated, 2.8% unknown ductility).

Based on these observation it is decided to limit the range of buildings considered here to non-ductile RC frame buildings without seismic design, which constitute about 89.5% of the sampled buildings. Additional information on the structural parameters, such as material strength, typical dimensions of the component elements, and other parameters (see Table 5.4), which cannot be obtained from rapid visual screening, are available from Grigoratos et al. (2016) and Al-Dabbeek and Al-Jawhari (2002). These can be used to subdivide the non-ductile RC frame buildings into sub classes according to their height and age, where the latter is based on changes in building practice regarding the strength of used materials (Table 5.4). In total, 16 sub-types of RC frame buildings without seismic design are defined using the same number of floors as Grigoratos et al. (2016), i.e., 2, 4, 7, and 10, but additionally distinguishing buildings built before 1980, the 1980s, 1990s and since 2000. A large portion (88%) of the buildings show irregularities, 56% have torsional eccentricity, 19.7% large overhangs or soft stories. As only the latter can be captured using single degree of freedom oscillators irregularities, it was decided to not consider irregularity within this thesis. The rich data set collected could be used in the future to define more realistic building types, but this is beyond the demonstration purpose here.

Mapping the observations of the survey to the defined categories yields a building type distribution for Nablus as shown in the $nr[\%]$ -column in Table 5.5.

Table 5.4.: Building parameter ranges from different sources and the employed values.

Building	Parameter	Range	Employed	
all	ground floor ¹	3-6 m	3.5 m	
	floors ¹	3-3.5 m	3.25 m	
	bay width ¹	4-8 m	split length until bays in range	
	beam spacing ³	0.5 m	0.5 m	
	beam width ³	0.1 m	0.1 m	
	concrete cover		2 cm	
	concrete yield strain		0.002	
	concrete ult. strain		0.0035	
	concrete weight density		25 kN/m ³ (EC-1)	
	steel elasticity		204'500 MPa	
	ult. steel strength ²		72 + 1.3f _{s,y} MPa	
	lap splice length		25% floor height	
	1/2 way ribbed slab	slab depth ¹	0.2-0.5m	0.35 m
		hidden beam depth ¹	0.25-0.5m	0.375 m
solid slab	slab depth ¹	0.15-0.3m	0.225 m	
	drop beam depth ¹	0.4-0.8m	0.6 m	
-1980	concrete strength ¹	15-20 MPa	17.5 MPa	
1980-1990		20-25 MPa	22.5 MPa	
1990-2000		28 MPa	28 MPa	
2000+		28-32 MPa	30 MPa	
-1990	steel strength ¹	370 MPa	370 MPa	
1990+		420 MPa	420 MPa	
-1980	stirrup diameter ¹	6-8 mm	7 mm	
1980+		10 mm	10 mm	
-1980	stirrup spacing ¹	25 cm	25 cm	
-1980		15-20 cm	17.5cm	

¹ Grigoratos et al. (2016)

² least squares linear model using PEER Structural Performance Database

³ Al-Dabbeek and Al-Jawhari (2002)

Table 5.5.: Building model capacity, price (\pm standard deviation) and share in numbers and value. The model-id is composed by the types typical number of floors and the bounds for the year of their construction.

model-id	Capacity [MN]/[%]		Ln. price [USD]		nr[%]	value [%]
	V_y/d_y	V_u/d_u	μ	σ		
2f 1980	0.73±0.22/1.53±0.47	0.59±0.16/3.43±0.74	12.42	0.72	7.2	9.7
2f 1980-1990	0.67±0.03/0.71±0.18	0.53±0.03/3.13±0.63	12.32	0.62	6.3	6.6
2f 1990-2000	1.53±0.30/0.77±0.07	1.26±0.27/3.54±0.51	12.39	0.61	6.3	5.1
2f 2000	0.62±0.30/1.32±0.38	0.51±0.23/4.35±0.89	12.57	0.54	8.2	7.4
4f 1980	0.82±0.05/1.30±0.25	0.68±0.03/2.56±0.50	13.26	0.62	11.6	13.7
4f 1980-1990	1.27±0.37/0.99±0.16	1.03±0.30/4.49±1.74	13.24	0.69	13.8	15.5
4f 1990-2000	1.74±0.19/1.01±0.18	1.39±0.18/3.28±0.51	13.45	0.53	11.4	9.7
4f 2000	1.11±0.42/1.79±0.42	0.89±0.32/4.75±1.24	13.34	0.56	14.4	13.7
7f 1980	2.14±0.44/2.00±0.69	1.95±0.55/8.78±4.23	14.47	0.69	0.4	0.5
7f 1980-1990	2.11±0.47/1.15±0.08	1.71±0.39/4.68±1.31	14.10	0.29	1.6	1.2
7f 1990-2000	2.28±0.10/1.09±0.08	1.78±0.15/4.88±0.93	14.21	0.66	3.3	2.8
7f 2000	2.31±0.07/2.43±0.50	1.89±0.08/9.92±2.59	14.17	0.52	5.7	5.4
10f 1980	2.46±0.56/1.46±0.09	2.02±0.49/7.10±3.53	14.79	0.01	1.3	1.9
10f 1980-1990	2.98±0.42/1.15±0.03	2.43±0.35/5.91±0.09	14.28	0.58	2.2	1.3
10f 1990-2000	2.88±0.44/1.12±0.09	2.37±0.34/6.58±0.97	14.61	0.48	1.6	2.0
10f 2000	3.12±0.66/2.09±0.50	2.66±0.72/9.22±3.31	14.88	0.56	4.7	3.5

The proportion of concrete buildings is estimated at 95% of the total number of buildings, considering only the median (12'639), a single concrete exposure set (12'012 buildings) is generated composed of the 16 building types with a distribution as listed in Table 5.5. Spatial variability within the town can be considered

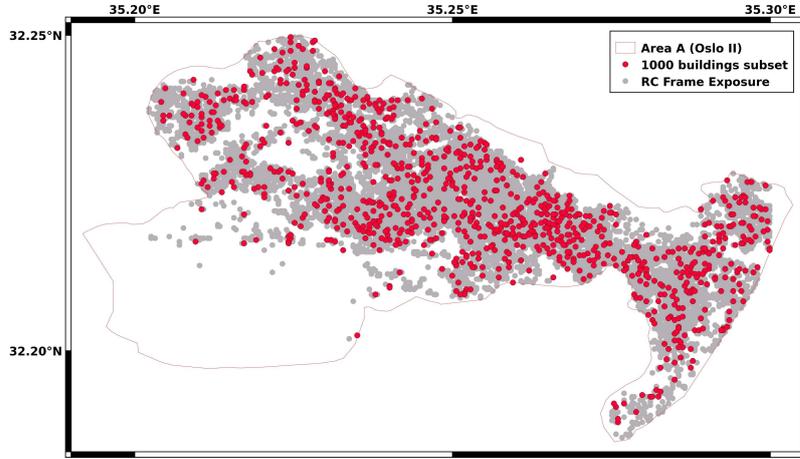


Figure 5.14.: Spatially distributed exposure model of RC frame structures in Nablus (gray points). For the risk analysis, only a random sample of 1'000 buildings (red) was used (see Figure 5.9).

using the strata identified from the satellite analysis. However, here only the identified built-up region was used to randomly distribute the buildings within the town using a spatial probability distribution derived from the road density. The road density was estimated by converting streets available from OSM into 10 m point chains and computing point density on a grid discretized with a 100 m resolution. The resulting spatial distribution is shown in Figure 5.14 for the subsequent risk analysis, however, only a subset of 1'000 randomly sampled buildings (red points) will be used. For the subsequent fragility analysis soil types are distinguished according to Eurocode-8 (European Committee for Standardization (2004b), see Table C.1). Based on the v_{s30} values from topography (see Figure 5.5), only soil type B ($v_{s30} = 360 - 800m/s$) and C ($v_{s30} = 180 - 360m/s$) are found in Nablus.

The resulting building type distribution, including soil type, is shown in Table 5.6 for the full model and the sample.

Table 5.6.: Building type with Eurocode EC-8 soil type derived from v_{s30} , and their proportion p [%] in the *full* exposure model and the *sample* used for risk estimates herein (red points in Figure 5.14). Note the missing values are not defined in the respective model.

building type	p_{full} [%]	p_{sample} [%]
2f 1980 B	7.15	8.80
2f 1980 C	0.02	0.10
2f 1980-1990 B	6.35	6.90
2f 1980-1990 C	0.04	
2f 1990-2000 B	6.24	5.20
2f 1990-2000 C	0.03	0.10
2f 2000 B	8.13	7.80
2f 2000 C	0.09	0.10
4f 1980 B	11.49	12.40
4f 1980 C	0.13	0.3
4f 1980-1990 B	13.74	14.60
4f 1980-1990 C	0.09	
4f 1990-2000 B	11.3	10.90
4f 1990-2000 C	0.06	
4f 2000 B	14.3	13.40
4f 2000 C	0.06	
7f 1980 B	0.38	0.50
7f 1980 C	0.01	
7f 1980-1990 B	1.57	1.40
7f 1980-1990 C		
7f 1990-2000 B	3.26	2.70
7f 1990-2000 C	0.01	0.10
7f 2000 B	5.7	5.40
7f 2000 C	0.04	
10f 1980 B	1.28	1.80
10f 1980 C	0.02	
10f 1980-1990 B	2.21	1.50
10f 1980-1990 C	0.02	
10f 1990-2000 B	1.55	1.80
10f 1990-2000 C	0.02	0.10
10f 2000 B	4.7	4.10
10f 2000 C	0.01	

5.3. Fragility models for Nablus

In order to derive state-dependent fragility models for the defined building types the simplified column based approach as described in section 4.2. Fundamentally, the approach is to approximate the buildings by bare moment frames without any seismic design using the few known building parameters (see Table 5.4). As it is unknown, which building code (if any) was used to design the buildings in Nablus, gravity load design prescriptions for concrete buildings as described by Eurocode-2 (European Committee for Standardization 2004a) are used (see Appendix C.1).

For the resulting building models non-linear static pushover curves are determined using EC-8 and ESDOF systems are obtained through the Rayleigh method. Capacities (Table 5.5) resulting from this oversimplified approach are unrealistically high. Considering deterioration of the ESDOF through secant stiffness and empirical damping and running the IDA on these with a set of ground motion records chosen for Nablus according to EC-8 yields fragilities models as listed in Table 5.7. Since no building parameter variations except slab type (see Table 5.5) were considered, only a small epistemic is encoded in the fragility curves (indicated by $\pm\sigma$). In accordance with the current limitation of the implementation, the fragilities are expressed in PGA. As expected from the strong capacities, the usage of PGA, the simplified deterioration and only a linear system, derived fragilities are very low (Table 5.7).

Figure 5.15a shows fragilities for an intact 7 story building built in Nablus after 2000, compared to moment frame fragility models (with various settings) available from the OpenQuake platform. Especially higher damage states are not captured realistically, where median and standard deviations are very high. One can conclude from this that any subsequent risk estimate will be strongly underestimated. As Figure 5.15d shows, a state dependent effect can be modeled (here only exceedance probabilities for ds_4 , are shown starting from different damage states). The effect is not pronounced and likely to be an underestimation as well, however the models should be sufficient to demonstrate if state dependency is relevant for urban seismic risk assessments. In an actual risk assessment, more care should be taken for the development of fragility models, at the least adopting the simplified approach as suggested by Bazzurro et al. (2004). As no recent damaging events occurred, the initial damage states for the buildings in Nablus has been set to an intact state.

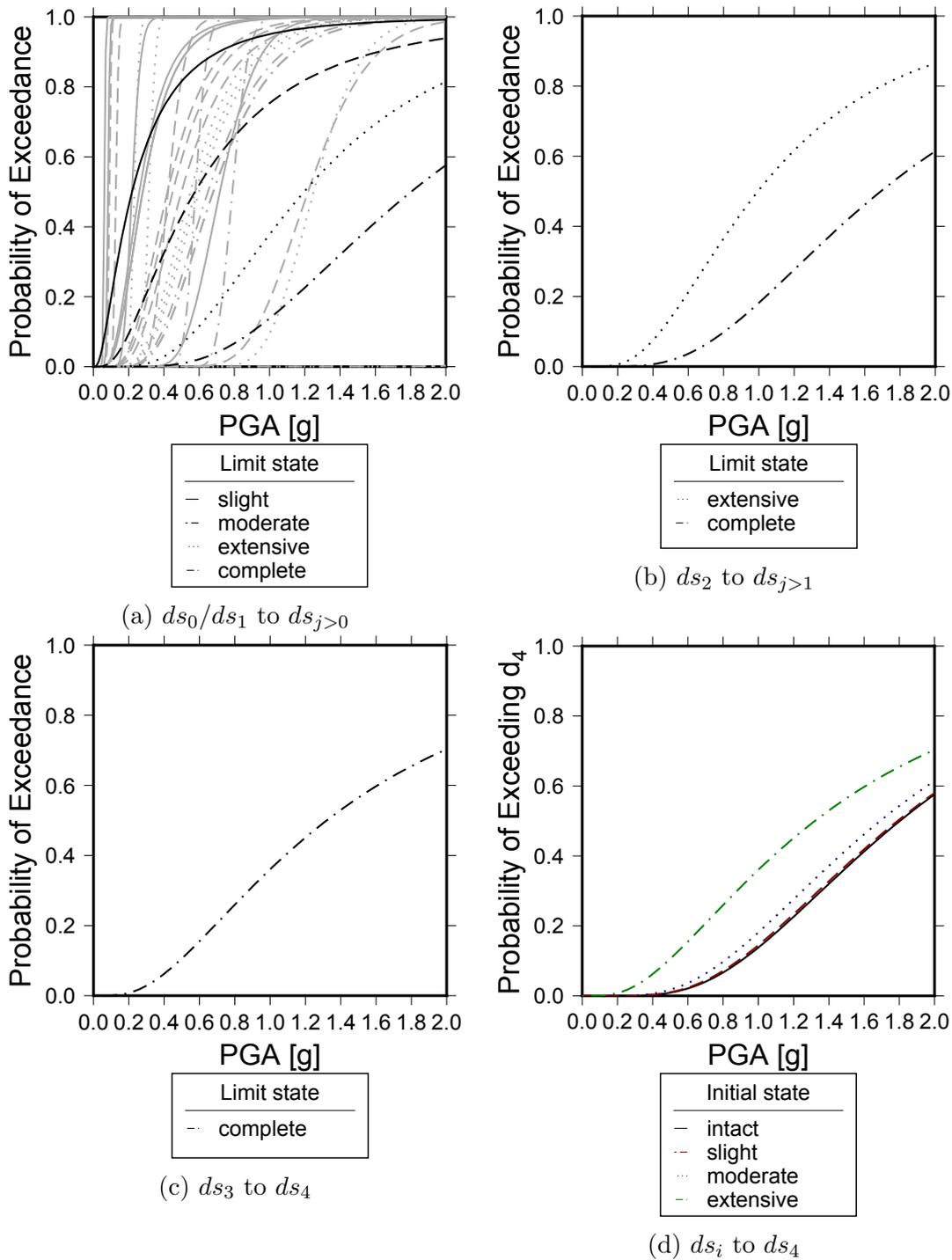


Figure 5.15.: Fragility curves for a 7 storey RC frame built after 2000 with a one-way ribbed slab on EC-8 soil type B, for 5.15a-5.15c different starting damage states (ds_i). Note that grey lines in 5.15a are literature models as obtained from the OpenQuake database (available at <https://platform.openquake.org>; accessed on May 2018). 5.15a shows curves for ds_0 and ds_1 as ds_1 is still in the elastic range they are identical. 5.15d shows exceedance probabilities for drifts corresponding to ds_4 from the different starting damage states.

Table 5.7.: Lognormal fragility curves for building models on soil type B and C (\pm standard deviation). Only simulated fragilities are shown while the others are calculated using equation 4.20.

model-id	F_{ij}	μ_B	σ_B	μ_C	σ_C
10f 1980	F01	-1.997 \pm 0.002	0.46 \pm 0.005	-1.977 \pm 0.003	0.691 \pm 0.005
	F12	-1.234 \pm 0.002	0.46 \pm 0.005	-1.214 \pm 0.003	0.691 \pm 0.005
	F23	-0.379 \pm 0.012	0.421 \pm 0.002	-0.31 \pm 0.016	0.629 \pm 0.005
	F34	0.237 \pm 0.016	0.39 \pm 0.001	0.348 \pm 0.021	0.587 \pm 0.003
10f 1980-1990	F01	-2.204 \pm 0.184	0.75 \pm 0.003	-2.478 \pm 0.185	0.9 \pm 0.006
	F12	-1.442 \pm 0.184	0.75 \pm 0.003	-1.716 \pm 0.185	0.9 \pm 0.006
	F23	-0.563 \pm 0.044	0.721 \pm 0.001	-0.844 \pm 0.042	0.83 \pm 0.006
	F34	0.053 \pm 0.021	0.672 \pm 0.006	-0.194 \pm 0.02	0.776 \pm 0.008
10f 1990-2000	F01	-1.805 \pm 0.03	0.763 \pm 0.001	-2.124 \pm 0.032	0.943 \pm 0.003
	F12	-1.043 \pm 0.03	0.763 \pm 0.001	-1.362 \pm 0.032	0.943 \pm 0.003
	F23	-0.397 \pm 0.048	0.695 \pm 0.001	-0.688 \pm 0.048	0.876 \pm 0.005
	F34	0.158 \pm 0.055	0.662 \pm 0.002	-0.116 \pm 0.054	0.832 \pm 0.004
10f 2000	F01	-1.722 \pm 0.028	0.761 \pm 0.001	-2.029 \pm 0.029	0.953 \pm 0.0
	F12	-0.96 \pm 0.028	0.761 \pm 0.001	-1.267 \pm 0.029	0.953 \pm 0.0
	F23	-0.37 \pm 0.007	0.69 \pm 0.001	-0.66 \pm 0.006	0.898 \pm 0.002
	F34	0.163 \pm 0.003	0.658 \pm 0.001	-0.118 \pm 0.004	0.86 \pm 0.004
2f 1980	F01	-1.579 \pm 0.141	0.612 \pm 0.228	-1.243 \pm 0.285	0.521 \pm 0.1
	F12	-0.817 \pm 0.141	0.612 \pm 0.228	-0.481 \pm 0.285	0.521 \pm 0.1
	F23	0.016 \pm 0.056	0.566 \pm 0.195	0.377 \pm 0.217	0.546 \pm 0.048
	F34	0.642 \pm 0.028	0.538 \pm 0.173	1.016 \pm 0.195	0.545 \pm 0.033
2f 1980-1990	F01	-1.458 \pm 0.162	0.352 \pm 0.007	-1.313 \pm 0.172	0.573 \pm 0.011
	F12	-0.696 \pm 0.161	0.352 \pm 0.007	-0.551 \pm 0.172	0.573 \pm 0.011
	F23	-0.034 \pm 0.004	0.35 \pm 0.008	0.143 \pm 0.005	0.564 \pm 0.012
	F34	0.537 \pm 0.071	0.363 \pm 0.009	0.709 \pm 0.069	0.56 \pm 0.01
2f 1990-2000	F01	-1.315 \pm 0.029	0.408 \pm 0.008	-1.173 \pm 0.052	0.636 \pm 0.027
	F12	-0.551 \pm 0.03	0.408 \pm 0.008	-0.411 \pm 0.052	0.636 \pm 0.027
	F23	0.176 \pm 0.007	0.4 \pm 0.006	0.358 \pm 0.018	0.552 \pm 0.01
	F34	0.768 \pm 0.003	0.39 \pm 0.003	0.947 \pm 0.005	0.525 \pm 0.005
2f 2000	F01	-1.275 \pm 0.01	0.443 \pm 0.023	-1.169 \pm 0.017	0.681 \pm 0.005
	F12	-0.513 \pm 0.01	0.443 \pm 0.023	-0.407 \pm 0.017	0.681 \pm 0.005
	F23	0.177 \pm 0.006	0.415 \pm 0.006	0.334 \pm 0.016	0.576 \pm 0.005
	F34	0.75 \pm 0.007	0.399 \pm 0.003	0.922 \pm 0.01	0.542 \pm 0.008
4f 1980	F01	-1.705 \pm 0.002	0.39 \pm 0.001	-1.545 \pm 0.002	0.575 \pm 0.001
	F12	-0.943 \pm 0.002	0.39 \pm 0.001	-0.783 \pm 0.002	0.575 \pm 0.001
	F23	-0.153 \pm 0.024	0.383 \pm 0.001	0.05 \pm 0.024	0.527 \pm 0.0
	F34	0.468 \pm 0.033	0.38 \pm 0.0	0.645 \pm 0.032	0.529 \pm 0.001
4f 1980-1990	F01	-1.562 \pm 0.012	0.432 \pm 0.002	-1.452 \pm 0.011	0.684 \pm 0.001
	F12	-0.8 \pm 0.012	0.432 \pm 0.002	-0.69 \pm 0.011	0.684 \pm 0.001
	F23	-0.125 \pm 0.044	0.413 \pm 0.001	0.037 \pm 0.044	0.575 \pm 0.001
	F34	0.443 \pm 0.061	0.398 \pm 0.001	0.617 \pm 0.062	0.539 \pm 0.002
4f 1990-2000	F01	-1.376 \pm 0.058	0.47 \pm 0.0	-1.345 \pm 0.059	0.678 \pm 0.001
	F12	-0.614 \pm 0.058	0.47 \pm 0.0	-0.582 \pm 0.059	0.678 \pm 0.001
	F23	0.006 \pm 0.037	0.424 \pm 0.0	0.086 \pm 0.036	0.618 \pm 0.001
	F34	0.547 \pm 0.069	0.4 \pm 0.003	0.658 \pm 0.074	0.588 \pm 0.003
4f 2000	F01	-1.218 \pm 0.006	0.441 \pm 0.001	-1.206 \pm 0.007	0.701 \pm 0.0
	F12	-0.456 \pm 0.006	0.441 \pm 0.001	-0.444 \pm 0.007	0.701 \pm 0.0
	F23	0.055 \pm 0.004	0.416 \pm 0.0	0.115 \pm 0.005	0.64 \pm 0.001
	F34	0.553 \pm 0.004	0.396 \pm 0.0	0.637 \pm 0.004	0.611 \pm 0.001
7f 1980	F01	-1.619 \pm 0.011	0.459 \pm 0.012	-1.521 \pm 0.013	0.682 \pm 0.003
	F12	-0.857 \pm 0.011	0.459 \pm 0.012	-0.759 \pm 0.013	0.682 \pm 0.003
	F23	-0.155 \pm 0.015	0.42 \pm 0.002	-0.004 \pm 0.014	0.578 \pm 0.001
	F34	0.42 \pm 0.024	0.401 \pm 0.001	0.591 \pm 0.024	0.547 \pm 0.003
7f 1980-1990	F01	-1.55 \pm 0.041	0.475 \pm 0.006	-1.511 \pm 0.05	0.671 \pm 0.007
	F12	-0.788 \pm 0.041	0.475 \pm 0.006	-0.748 \pm 0.05	0.671 \pm 0.007
	F23	-0.151 \pm 0.028	0.426 \pm 0.001	-0.066 \pm 0.034	0.613 \pm 0.005
	F34	0.395 \pm 0.023	0.4 \pm 0.002	0.513 \pm 0.028	0.584 \pm 0.003
7f 1990-2000	F01	-1.701 \pm 0.003	0.751 \pm 0.003	-1.975 \pm 0.005	0.899 \pm 0.006
	F12	-0.939 \pm 0.003	0.751 \pm 0.003	-1.213 \pm 0.005	0.899 \pm 0.006
	F23	-0.21 \pm 0.11	0.721 \pm 0.001	-0.49 \pm 0.112	0.829 \pm 0.005
	F34	0.359 \pm 0.15	0.679 \pm 0.005	0.11 \pm 0.154	0.782 \pm 0.009
7f 2000	F01	-1.455 \pm 0.021	0.726 \pm 0.006	-1.758 \pm 0.009	0.919 \pm 0.003
	F12	-0.693 \pm 0.021	0.726 \pm 0.006	-0.996 \pm 0.009	0.919 \pm 0.003
	F23	-0.169 \pm 0.014	0.714 \pm 0.002	-0.464 \pm 0.017	0.846 \pm 0.003
	F34	0.327 \pm 0.025	0.684 \pm 0.0	0.057 \pm 0.029	0.808 \pm 0.005

5.4. Building value and loss ratios

Each simulation returns a chronological sequence of damage and repair events for each building over the defined observation period. Repeating the simulation in a Monte-Carlo framework, multiple independent damage histories are simulated. From these damage histories, one can calculate losses, considering loss as a fixed percentage of the buildings value, which occurs at the time of a damaging event and each transition from an initial damage state d_i to a damage state d_j as a separate loss event. For the application here only structural and non-structural losses are considered. Summing the Hazus MH structural, non-structural and drift-sensitive non-structural loss ratios for multi-family dwellings (Federal Emergency Management Agency 2011, Table 15-2, 15-3, 15-4) and converting the resulting cumulative damage loss function (2% for *slight*, 10% for *moderate*, 41.3% for *extensive* and 100% for *complete*) to an incremental loss ratio function.

Combined with building replacement costs, the loss ratios yield the nominal loss. Replacement costs usually correspond not to construction cost or the value of a building. This is due to factors such as, e.g., demand surge in case of a large scale catastrophe, or costs for demolishing and debris removal. Lacking data on replacement costs in Nablus, more readily available market prices are used as a proxy for replacement costs in the application here. The price of RC frame buildings in the West Bank was estimated by Al-Dabbeek and Al-Jawhari (2002) for 2002 as 60'000 USD for 200-300 m^2 housing units, with a range of 200-300 USD/ m^2 . More recent data from Numbeo (Feb 2016 - Aug 2017 20 values; www.numbeo.com as of August 2017) indicate prices between 1500 and 2500 ILS/ m^2 (422 to 705 USD/ m^2) for apartments in the city center. In comparison, Grigoratos et al. (2016) indicate replacement costs as 550-600 USD/ m^2 for an apartment, 750 USD/ m^2 for individual houses, and up to 900 USD/ m^2 for villas. Assuming that building values are age-independent and setting replacement costs to market price, a value is calculated for each building in the collected building data set, sampling the price range of 422 to 705 USD/ m^2 and multiplying by footprint size (OSM) and number of floors (RRVS). For each building type, these samples follow roughly a lognormal distribution (see fitted distributions in Table 5.5), which can be used to sample a value for each of the 12'012 RC frame buildings in the full exposure model. The sample of 1'000 buildings used here for demonstration purposes has a total value of 888 Mio USD.

With the assigned replacement costs and the defined loss ratios, the losses can be calculated from the damage and repair history for each building. Within the implemented framework, this is part of the post processing routine (see Figure 4.7).

6. Simulation results

6.1. Resulting loss estimates

Using the components defined for Nablus, independent simulations for a defined observation period can now be run. Converting the damages resulting from the simulations to their corresponding losses, the loss exceedance probabilities (EP) can be calculated, corresponding to the percentiles of the loss distribution. These give the probability to exceed a certain amount of loss within the observation period. As sequences might span several years, a 50 years period is chosen to clearly highlight the effects. Defining a target resolution of 1% probability of exceedance (PoE) for loss within this period, requires a minimum of 100 simulations. In order to obtain stable results, it is assumed that at least 10 simulations (i.e., 500 simulated years) should define the loss at 1% and thus at least 1'000 simulations are necessary. If model uncertainties are to be determined, using the bootstrap approach increases the number of simulations required, depending on the sampling ratio. For example, with 10% sampling a minimum of 10'000 simulations must be run to reliably resolve 1% PoE.

Total absolute losses

The solid curve in Figure 6.1a shows a median EP-curve for 50 year losses considering full seismic sequences (**ETAS**), a *conservative* time delayed repair (**DR**, $\mu_{ln} = 1$, $\sigma_{ln} = 0.5$), and state-dependent fragility (**SD**), aggregated for all 1'000 buildings defined in the exposure sample. Following the previous reasoning, the curve is calculated from a distribution of EP-curves resulting from 1'000 independent bootstraps, where each bootstrap is composed of a random sample of 10% of all 10'000 simulations, i.e., 1000 simulations. The support of an EP-curve for a single bootstrap sample increases with the exceedance probability, i.e., it ranges from 10 simulations at 1% probability of exceedance (PoE), until 100 simulations for 10% PoE corresponding to 500 and 5000 simulation years, respectively. Due to the underestimation of building fragility the absolute loss values shown here are most likely an underestimate of the actual risk. In addition to the median, Figure 6.1a shows quantiles of the distribution resulting from the bootstraps corresponding to the 5 and 95 (dotted lines) quantiles, and the 25 and 75 % quantiles (dashed lines) of the loss distribution. While Figure 6.1a shows the proposed full model (*ETAS+DR(1,0.5)+SD*), Figure 6.1b shows the estimates resulting from the classical risk assumptions, i.e., Poissonian inter event time and immediate repair (fixed

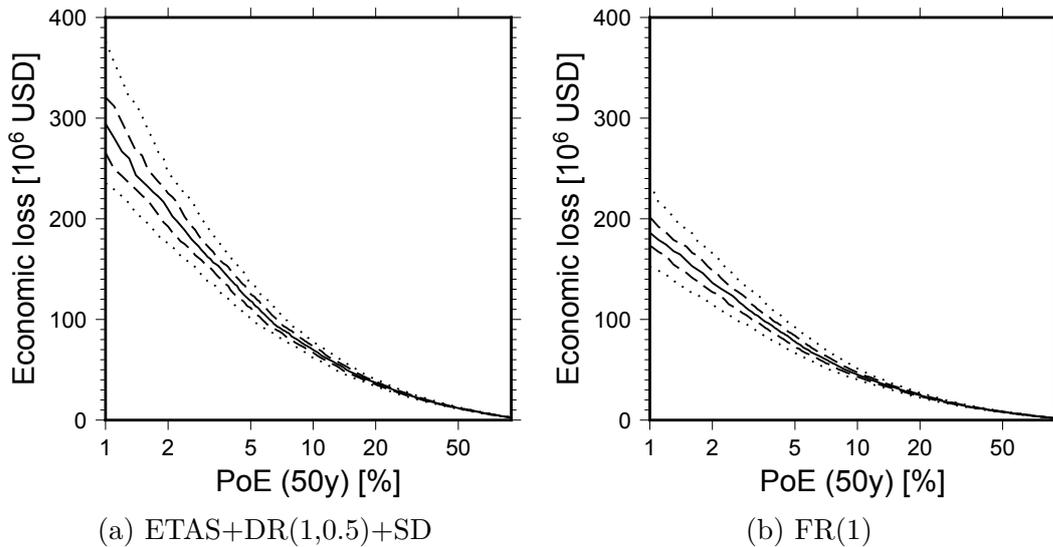


Figure 6.1.: 50 year median loss exceedance probability (PoE) curve for 1'000 simulated buildings in Nablus, determined using a 10% bootstrap sampling (1'000 samples) on 10'000 simulations alongside with the 5, 25, 75 and 95% quantiles (dotted, dashed, dashed, and dotted, respectively). Left side shows curves resulting from using a model with full seismicity (ETAS), delayed repair (DR; $\ln(\mu) = 1, \ln(\sigma) = 0.5$ [yrs]), and state dependent fragility (SD). The right side shows a model using a classical assumption (Poissonian seismicity) and fixed probability (FR(1)), i.e., immediate repair.

repair **FR** with a probability of 1) of damaged buildings.

While the hazard in the simulations considering full seismic sequences increased by 20.5% for the 1% probability of exceedance value in 50 years ($0.53g$ instead of $0.44g$ at $v_{s30} = 760m/s$; see Figure 5.6 in Section 5.1.4), the median loss for a time- and state-dependent model increases by about 58.4% at 1% PoE with respect to a classical risk model (293 vs. 185 Mio USD). Considering the total value of the buildings in the sample exposure model of 888 Mio USD, this corresponds to approximately 33% and 21% of the total value, respectively. At 1% PoE, the deviation from the median is about 68 Mio USD for the 5 and 95% quantiles of the *ETAS+DR(1,0.5)+SD* model's loss distribution and about 40 Mio for the *FR(1)* model, which corresponds to $\pm 23\%$ and $\pm 22\%$ variations, respectively. With an increasing number of supporting simulations, i.e., at higher PoEs, this shrinks to about $\pm 11\%$ and $\pm 12\%$, respectively at 10% PoE. Note that aleatory uncertainty is almost exclusively included here, as epistemic uncertainty is only considered for the fragility models to a small extent (3 different slab types; see Table 5.5 and $\pm\sigma$ in Table 5.7).

Figure 6.2 gives the residuals between both models, revealing a strong exponential increase in the difference with lower PoEs (note the logarithmic x-axis). While the residual at 20% PoE is about 12 Mio USD, it increases to 25 Mio USD at 10% PoE, 41 Mio USD at 5%-PoE, until it reaches 108 Mio USD difference at 1% PoE.

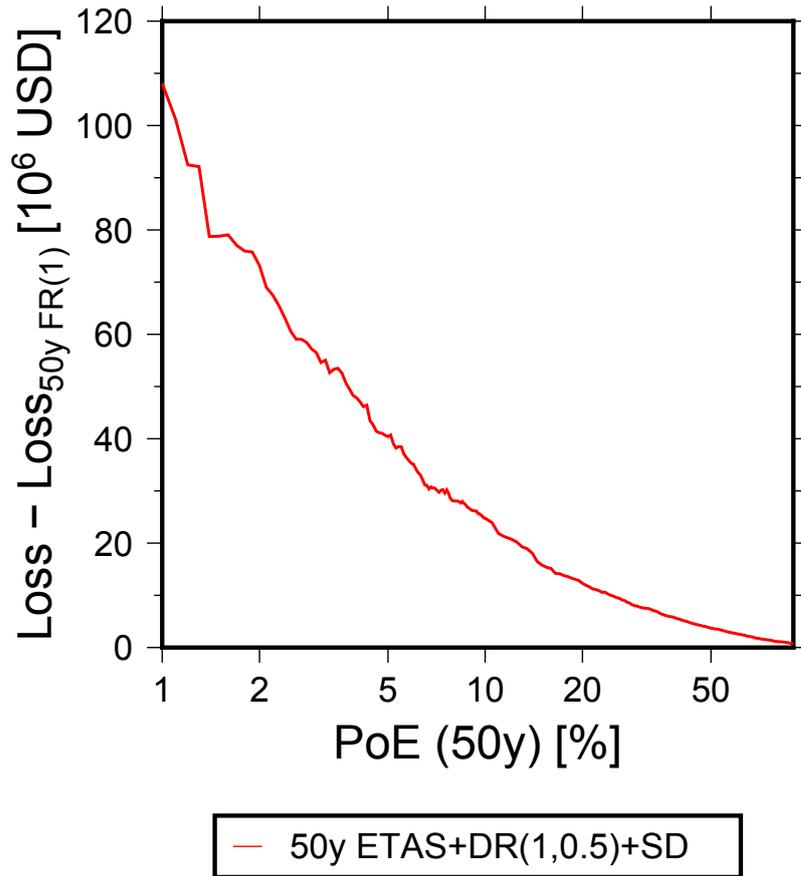


Figure 6.2.: Residual between the median loss EP-curves of Figure 6.1 (calculated as median EP-curve Figure 6.1a minus median EP-curve Figure 6.1b). Note the logarithmic scale for the probability.

Aggregating losses per building type

While the previous results were aggregated over all buildings and yield an overview result for all assets, aggregating losses per building type can give insight into each type's contribution to these total losses. Figure 6.3 shows two median EP curves aggregated per building type and associated percentiles. While 6.3a shows curves for 4 story buildings built until 1980, 6.3b shows 10 story buildings built in the 80s, with both types on soil type B for the *ETAS+DR(1,0.5)+SD* model. The first represents the building type contributing most to the total losses (80 Mio USD at 1% PoE), the latter is on the lower end of the contributors (7 Mio USD at 1% PoE). There are several building types whose losses in terms of the total loss are negligible, when compared to the other building types (< 1 Mio USD at 1% PoE). These are on soil type C (2f: -1980, 1990-2000, 2000; 7f: 1990-2000; 10f: 1990-2000; see Table 5.6) and of little relevance due to their small number of only 1 building each in the sample. An exception among soil type C buildings are 4 story buildings built till 1980, representing 0.3% of the buildings, they contribute 7 Mio USD at 1%, corresponding to 2.4% of the total loss. On soil type B, only 7 story buildings built before 1980, representing 0.5% of the sample, have a negligible contribution to the total losses (1 Mio at 1% PoE). All other types range between the two building types shown in Figure 6.3. Please note that the 1'000 sampled buildings do not contain

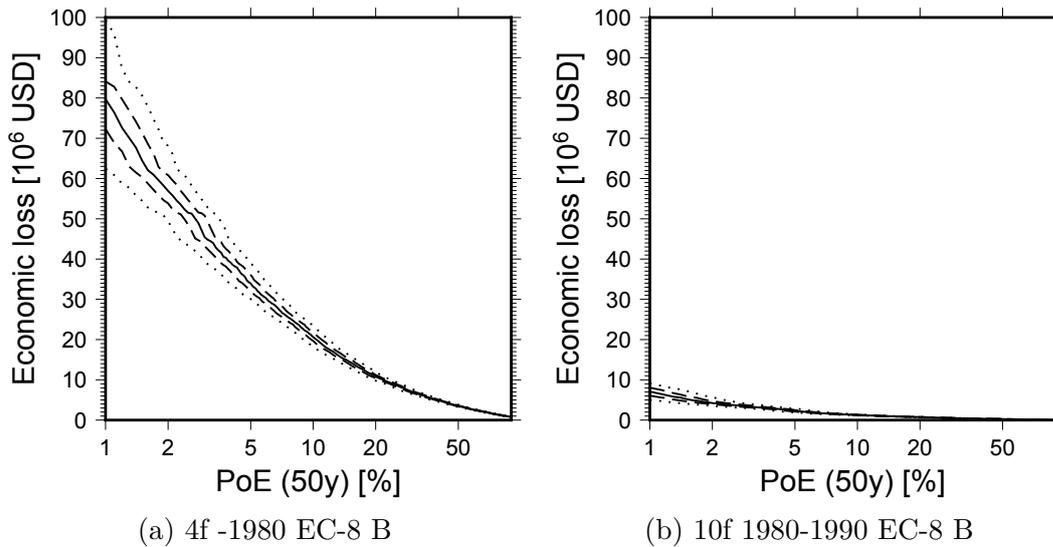


Figure 6.3.: 50 year median loss exceedance probability (PoE) curves from the time- and state-dependent model (Figure 6.1a aggregated for two individual building types, using a 10% bootstrap sampling approach and 5, 25, 75, and 95% percentiles (dotted, dashed, dashed, and dotted, respectively). 6.3a shows the building type that contributed the most, while 6.3b shows one of the smallest (non-negligible, i.e., > 1 Mio at 1% PoE) contributions.

all potential building types on EC-8 soil class C, as only a few areas in Nablus are of this soil type (see Table 5.6 and Figure 5.5).

Loss/Value ratios per building type

The building type contributions to the total absolute losses does not allow one to judge the vulnerability of the individual building types directly as their numbers and resulting values are different. Figure 6.4 thus shows the median loss ratios for all building types. The loss ratios are obtained by dividing the median loss exceedance curve aggregated per building type by the total value of the respective type in the exposure model. These allow to identify more clearly the building types that are prone to causing high levels of losses under the simulated conditions. The left side of Figure 6.4 shows building types on soil type B (dark colors), whereas the right side shows those on type C (bright colors). Consistently with total loss contributors, the highest relative losses for soil type B (left plot) are expected from four storey buildings built until the 80s (dashed red). The least vulnerable buildings according to the simulations here are 7 storey buildings built since 2000 (dotted black) and 2 storey buildings from the 90s (solid green). On soil type C (right side plot), as for soil type B, the largest relative losses are simulated for 4 storey buildings built until 1980, followed by buildings from the 90s with seven (dotted green) and ten storeys (dot dashed green) buildings. The lowest vulnerability on C is simulated for 2 storey buildings from the 90s (ranked second on soil type B). Comparing the left and right side, the analysis suggests that the (fewer) buildings on soil type C show

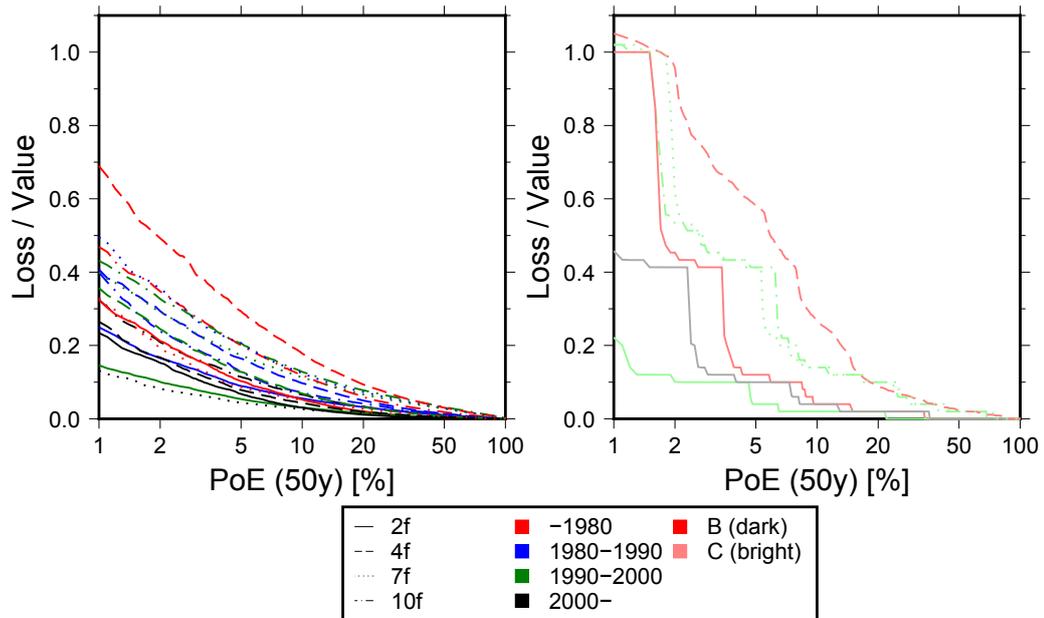


Figure 6.4.: Ratio of median loss EP-curves for $ETAS+DR(1,0.5)+SD$ (Figure 6.1a) and value of building stock, aggregated per building type in the subset of 1000 buildings in Nablus, again determined using 10% bootstraps. Left buildings on soil-type B (dark colors) and right on soil-type C (bright colors). Note that not all building types are encountered on soil type C (bright color tones) within this sample.

a tendency to higher vulnerability with respect to similar structures on soil type B. As expected, due to inferior materials, older structures built before 1980 (red) and to some extent also buildings built in the 80s (blue) tend to cause higher relative losses with respect to newer structures (black/gray and green). Note that the kinks in some of the curves are due to undersampling as buildings of this type are few in the sample.

6.2. Influence of model components on results

As mentioned earlier, the implementation of the simulation framework is able to consider the different aspects of the time- and state-dependent model separately. Rerunning the simulation with various settings, models are obtained that when compared to each other, allows one to judge the relevance of the individual simulation components within the presented framework. Note that pseudo random numbers were used in the stochastic models and in the sampling, both intended to ensure that the observed differences are not artificial. The analysis will be limited to the following factors: (1) Fore- and aftershocks (ETAS), (2) delayed repair (DR), and (3) state dependent fragility (SD). In addition, the (4) shape of the repair model in point (2) will be evaluated. Table 6.1 gives an overview over these models.

A first comparison can be made on the median estimates resulting from the bootstrap sampling for each simulation. Figure 6.5 compares the first six models of Table

Table 6.1.: Simulations with 50 year observation periods run with the developed framework under different modelling conditions.

ID	$T_{obs}[a]$ ¹	ETAS	Repair ²	State dependency
50y FR(1)	50	✗	FR(1)	✗
50y DR(1,0.5)	50	✗	DR(1,0.5)	✗
50y DR(1,0.5)+SD	50	✗	DR(1,0.5)	✓
50y ETAS+FR(1)	50	✓	FR(1)	✗
50y ETAS+DR(1,0.5)	50	✓	DR(1,0.5)	✗
50y ETAS+DR(1,0.5)+SD	50	✓	DR(1,0.5)	✓
50y ETAS+DR(-0.1,0.4)+SD	50	✓	DR(-0.1,0.4)	✓
50y ETAS+DR(1.3,0.5)+SD	50	✓	DR(1.3,0.5)	✓

¹ Observation period, i.e., simulated period after input catalog.

² DR: time delayed repair with underlying normal distribution of lognormal model (μ, σ); FR(p): fixed daily repair probability=p.

6.1, i.e., excluding the models with varying delayed repair models. Each of the median EP-curves is plotted relative to a model with classical risk assumptions, i.e., no ETAS simulations and a fixed repair probability of $p=1$, which corresponds to immediate repair (denoted as $FR(1)$). Every point of a model's EP-curve above a ratio $Loss/Loss_{FR(1)} = 1$ is thus higher than the reference model ($FR(1)$), and the contrary for points below it.

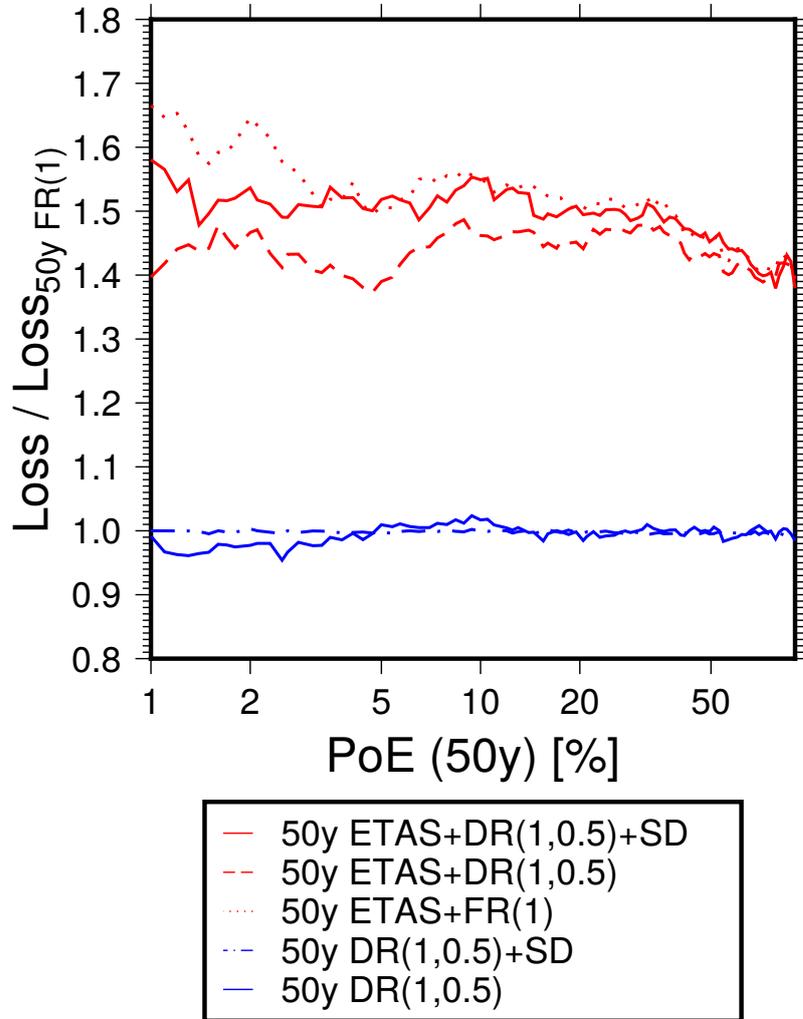


Figure 6.5.: Median loss exceedance probability curves, relative to the classical risk model (Figure 6.1b; $FR(1)$ in Table 6.1; relative loss equal to 1). Shown are curves for the first six models listed in Table 6.1 with varying modeling conditions such as modeled fore- and aftershocks (*ETAS*), time-delayed repair with log normal model ($DR(\ln(\mu), \ln(\sigma))$) with mean and standard deviation μ and σ , or repair with fixed probability p ($FR(p)$), and state-dependent fragility (*SD*).

ETAS vs Poissonian seismicity

The red and blue colors in Figure 6.5 distinguish models with full seismic sequences (*ETAS*) and models with Poissonian seismicity, respectively. The first observation from this comparison is that the *ETAS* models consistently yield higher losses over the full range of PoE than Poissonian models. This difference increases with decreasing exceedance probability, from about 40% at 1% PoE until up to 65% at 1% PoE. While all *ETAS* models coincide at high PoEs, they separate when considering the lower probabilities.

The time and state-dependent model proposed in the previous sections, i.e., the *ETAS* model with conservative time-delayed repair and state dependent fragility (*ETAS+DR(1,0.5)+SD*; red solid line in Figure 6.5), shows losses that are greater by 40% compared to the *FR(1)* model at high PoEs, quickly increasing to about 50% higher losses at 30% PoE, and then slowly reaching the largest increase of about 58% at 1%. In general, the trend is fluctuating by an order of 2-5% loss increase along the whole range of the PoE.

Only ETAS or delayed repair

The models behind the dotted and dashed red lines in Figure 6.5 investigate, if considering only one component is sufficient to obtain results that are comparable to the more complex full model *ETAS+DR(1,0.5)+SD*. In both cases full seismic sequences are considered, but one maintains the immediate repair assumption (*ETAS+FR(1)*), and the other considers delayed repair, but no state-dependent fragility *ETAS+DR(1,0.5)*. According to the curves presented in Figure 6.5 neglecting either component still captures the general tendencies, but fails to yield accurate results, especially below 10% PoE. They consistently tend to yield either lower or higher losses when compared to the full model.

Dropping state dependency (*ETAS+DR(1,0.5)*; dashed red curve), the losses start to deviate from those of the full model (solid red curve) already at high PoEs ($\sim 60\%$), yielding a loss increase with respect to the classical reference model *FR(1)*, which is between 0 and 28% lower than the one observed for the full model. A long period oscillation can be observed in the relative losses, with local maxima at 30%, 10%, and 2% PoE and local minima at 5% and 20% PoE. These replicate the behavior of the delayed repair model in combination with different event magnitudes and triggered sequences, dominating different PoE ranges. In the model without state dependency, the repair model is damping the increase in losses. However, as the fragility is not increasing with the damage state, buildings don't get damaged from subsequent smaller ground motion events. Hence, when compared to a classical *FR(1)* model the increase in losses for the model without state-dependency is lower than that of the full model. This difference between the two models increases with lower PoEs as stronger events dominate these. Strong events are more likely to cause high damage states, which in turn show a pronounced increase in fragility when state-dependency is considered. This is not the case for the *ETAS+DR(1,0.5)*

Table 6.2.: Average number of transitions in 50'000 years for the different simulations (see Table 6.1); colors rank the number of transitions per damage grade from highest (dark red) to lowest (dark blue).

ID	ds_{01}	ds_{02}	ds_{03}	ds_{04}	ds_{12}	ds_{13}	ds_{14}	ds_{23}	ds_{24}	ds_{34}
FR(1)	313'568	47'597	12'336	6'484						
DR(1,0.5)	308'996	46'675	11'922	6'175	694	168	88	22	9	2
DR(1,0.5)+SD	308'427	46'871	12'138	6'379	1'294	208	104	96	18	15
ETAS+FR(1)	516'225	79'905	20'876	10'661						
ETAS+DR(1,0.5)	414'916	62'970	16'162	8'034	6'636	1'730	901	1'085	567	445
ETAS+DR(1,0.5)+SD	416'367	63'244	16'311	8'199	10'926	1'883	871	2'963	743	1'517
ETAS+DR(-0.1,0.4)+SD	424'099	64'200	16'568	8'260	9'433	1'628	751	2'717	678	1'358
ETAS+DR(1.3,0.5)+SD	411'139	62'492	16'031	7'986	11'474	2'057	899	3'076	766	1'585

model which employs fragilities that are independent of the current damage state and therefore the losses resulting from the model underestimate the risk.

The red dotted line investigates the case, if one only simulates full seismic sequences, but immediately repairs structures (*ETAS+FR(1)*). The model follows the full model (*ETAS+DR(1,0.5)+SD*) closely in the high exceedance probability range. Below 40% PoE, the model shows a tendency to yield higher losses with respect to the full model. Around 10% PoE, it temporarily drops below the state-dependent model, which is due to event sequences that yield moderate or severe damages in the first events and reach collapse due to increased fragilities in later events for the full model, but don't show in the model with immediate repair. With lower PoEs, the immediate repair model yields considerably larger losses than the full model. This becomes understandable when the individual damage transitions for the different simulations are compared in Table 6.2. Adding fore- and aftershocks, but immediately repairing damage after each event yields a behavior that suggests that the buildings are damaged and repaired multiple times during a sequence, even if the events are very close in time. For example, starting from an intact building, a main shock might see the building reach a moderate damage state (ds_3). For the full model with delayed repair (*ETAS+DR(1,0.5)+SD*), the building keeps the damage state until the time is sufficient to repair it. Ground motions of smaller events in the sequence following the main shock might not be sufficient to cause the building to reach a higher damage state, despite the increase in fragility due to state-dependency (*SD*). For the resulting losses, this means that a single loss event has occurred going from ds_0 to ds_3 . In contrast, for the immediate repair model (*ETAS+FR(1)*), if a building is going through the same sequence, it is repaired prior to each sequence, the building can thus reach smaller damage states in addition to the first transition, i.e., it goes from ds_0 to ds_3 as with the full model, but is then repaired immediately, now in a subsequent aftershock it may go from ds_0 to ds_1 , is repaired again, and might go from ds_0 to ds_1 again in another sufficiently strong aftershock. This yields an artificial double counting of loss events, and the total number of loss events are larger than for the full, time-delayed model. This can be seen from the, on average, higher number of transitions of smaller damage states in Table 6.2. In the beginning, both models agree because this effect is compensated by the increased fragilities due to considering state-dependency in the full model. However, at lower PoEs, as more and stronger sequences occur, the double counting effect surpasses the increased fragility effect. The difference becomes very large if going to PoEs lower than those shown in Figure 6.5, that yields a strong overestimation of the losses in the *ETAS+FR(1)* model case. This can be seen calculating EP-curves for

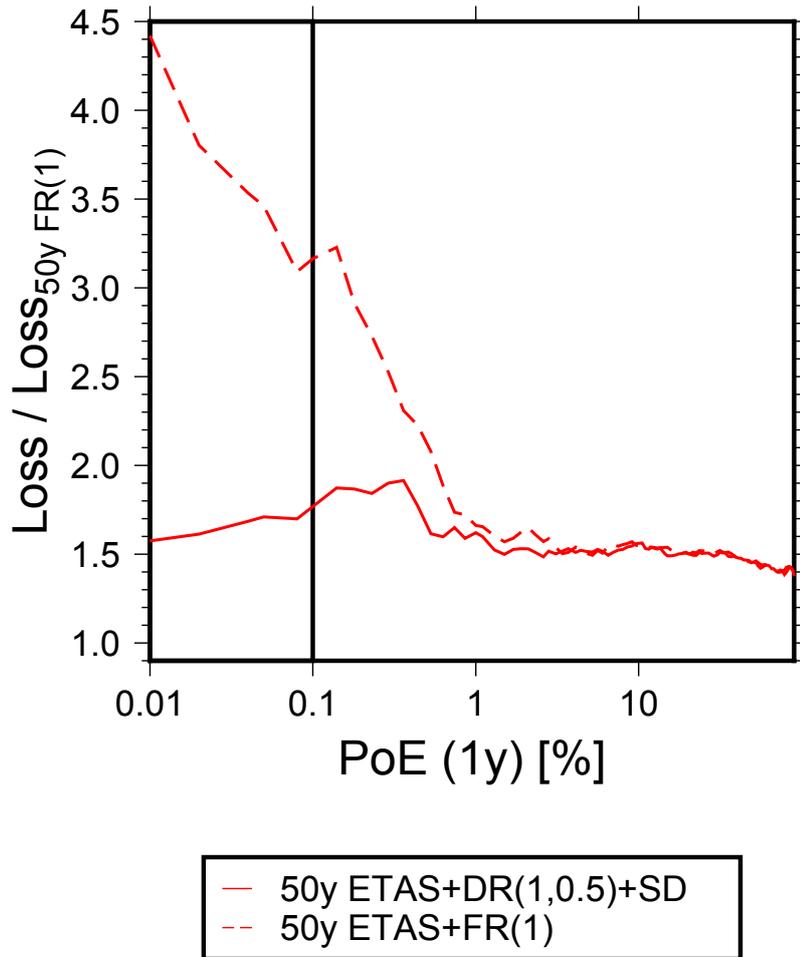


Figure 6.6.: EP curve for 1'000 simulated buildings in Nablus, without bootstrapping, calculated from 10'000 simulations for 50 years of seismicity. The gray marked area indicates the part of the curve that is supported by less than 500 simulated years.

both models without bootstrapping, which allows one to resolve lower PoEs, albeit at the cost of losing information on the model uncertainty. However, the strong trends shown in Figure 6.6 support the claim that the observed behavior is not an artifact of the sampling. Note that the gray marked area in Figure 6.6 indicates the highly uncertain part of the curve supported by less than 500 years of simulated seismicity.

Delayed repair or state-dependency in Poissonian model

Another important conclusion from Figure 6.5 can be drawn for the Poissonian models (blue curves), as they are deviating only slightly from the reference model $FR(1)$, i.e., $Loss/Loss_{FR(1)} = 1$. An enhanced view of these models is shown in Figure 6.7. The general observation from this comparison is that considering time-delayed repair (DR) and state dependency (SD) within a Poissonian seismicity model does not yield a pronounced increase in the risk estimates (note the different scale here). This observation strongly depends on the rate of occurrence of the strong ground

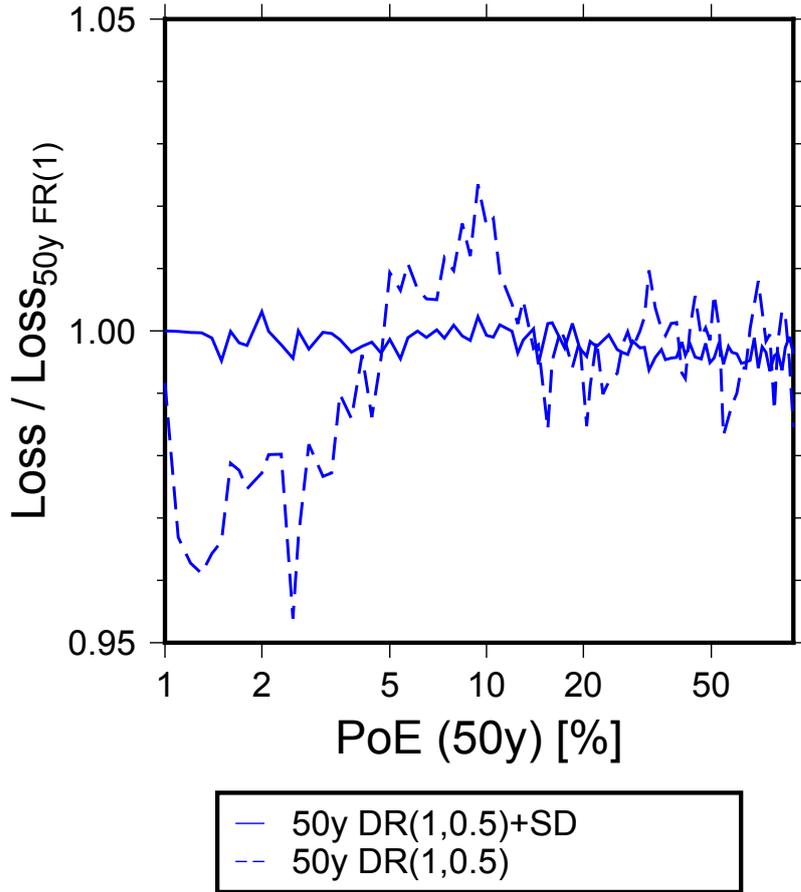


Figure 6.7.: Median loss EP-curves for the Poissonian models, one with only time delayed repair ($DR(1,0.5)$), and one with additional state dependent fragility $DR(1,0.5)+SD$, both relative to a classical risk model $FR(1)$.

motions and the chosen repair time model and might not be the case for other model conditions. However, despite the strong fragility models used here, considering that a relatively conservative repair time model was used, one can deduce with some confidence that a Poissonian model for Nablus does not benefit from such a model complication. Again, as already seen for the $ETAS+DR(1,0.5)$ results without state dependency, an oscillating behavior induced by the delayed repair model can be observed. The effect of damped losses is less strong as in the $ETAS$ equivalent, because fewer events (only Poissonian) are considered during the years causing the decrease. The state dependency counters this effect in the $DR(1,0.5)+SD$ model, yielding almost identical loss EP values as the reference model. It starts slightly below the classical reference model in the high probability range as delayed repair dominates for years with low activity, but then converges to $Loss/Loss_{FR(1)} = 1$ below 10% PoE. The small increase in relative loss between 5 and 10% is likely to be an artifact of the sampling. Unlike in the case when the two full seismic sequence models with fixed repair and delayed repair with state dependency were compared ($ETAS+DR(1,0.5)+SD$ vs $ETAS+FR(1)$), the fixed repair model does not yield higher losses at lower PoEs, because the time in between Poissonian events is sufficiently large to not show a detectable double counting effect.

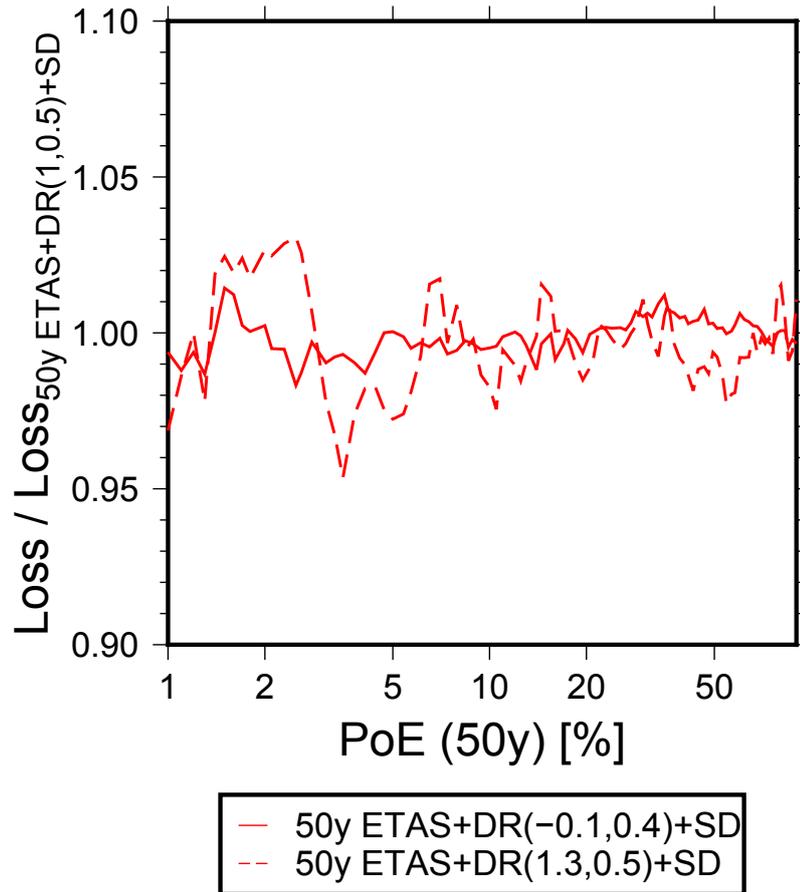


Figure 6.8.: Median loss exceedance probability curves for different time dependent repair models, relative to losses obtained from a model with the conservative repair model $ETAS+DR(1,0.5)+SD$. The solid curve shows losses for an optimistic model $ETAS+DR(-0.1,0.4)+SD$, the dashed line those of a pessimistic repair model $ETAS+DR(1.3,0.5)+SD$

Shape of the lognormal repair model

In the *ETAS* model, an influence should also be observed depending on the chosen probabilistic repair time model. As outlined in the repair model section (Figure 4.1) two alternative repair time models are proposed in addition to the *conservative* model ($DR(1,0.5)$) discussed so far. Figure 6.8 shows loss exceedance probabilities for two simulations considering full seismic sequences (*ETAS*), state dependency *SD*, and two different delayed repair models *DR*. One employs the *pessimistic* $DR(1.3,0.5)$ repair model and one the *optimistic* $DR(-0.1,0.4)$ model. Both are plotted relative to the model with the *conservative* delayed repair model. The figure shows that the behavior is complex and cannot be reduced to a simple rule where quicker repairs yield less loss, and the contrary. The observed variations are at the order of a few percent, not exceeding 5% difference and the trends are non-intuitive. For high PoEs ($>30\%$) the *optimistic* repair model (solid line) shows slightly ($\sim 1\%$) higher losses than the *conservative* model. Oscillating around a ratio of 1, the *optimistic* model slowly tends towards lower losses at the order of 1-2%. At about 1.2% PoE, observed losses are about 2% higher with respect to the *conservative* model.

The explanation for this can again be seen from the transitions listed in Table 6.2. While the *conservative* ($ETAS+DR(1,0.5)+SD$) model shows more transitions from higher damage states than the *optimistic* model ($ETAS+DR(-0.1,0.4)+SD$), it shows fewer transitions from the intact state. This is a similar effect as previously observed with the immediate repair model $FR(1)$. If the buildings are repaired during an ongoing sequence and not retrofitted accordingly, the losses might be higher than if the buildings are repaired at a later stage. On the other hand, the *pessimistic* model ($ETAS+DR(1.3,0.5)+SD$) shows a larger number of transitions to higher damage states, and lower numbers for transitions from the intact state. However, the corresponding dashed line in Figure 6.8 shows that the overall trend is towards lowering losses as buildings stay longer in high damage states and are not damaged again. In practice, the latter of course changes if not only transition losses are considered but also business interruption. At about 12, 6 and between 1.5 and 3 % PoE the losses for the *pessimistic* model are higher as the *conservative* model. This is due to the buildings staying in higher damage states, and the state dependency taking full effect. The oscillating behavior again observed between the different models reflects the different periods in the repair time models. Overall the influence of the model shape seems not very pronounced. Please, note that all these observations are limited as in practice only under rare circumstances buildings are reconstructed exactly as they were before. This is mainly the case for informal and uncoordinated reconstruction, and rarely also for monumental buildings. In most other cases changes in building practice or in seismic codes since the original construction of a building result in adjustments of the repaired or reconstructed building. Retrofitting or more resistant construction, if conducted properly, typically yields a less fragile building stock after the repair process is finished.

6.3. Statistical tests

In addition to the visual comparison, presented in the previous section, a brief statistical analysis is run on pair wise model combinations. The aim is to determine: 1) the statistical significance for observed differences between models, and 2) the practical relevance of these differences. It is assumed that a model ($i = 1, 2, \dots, m$) has a loss distribution at a certain level of exceedance probability p expressed as $f_i(x)$. The EP curves resulting from the bootstraps are considered samples X_{ij} of the actual loss distributions for model i along p , where $j = 1, 2, \dots, n$ and n is the number of bootstraps. If the underlying distributions at a level of exceedance probability, e.g. $p = 0.1$, for two models, say $i = 1$ and $i = 2$, are the same $f_1(x) = f_2(x)$, any observed differences in the mean values of the sample distributions, i.e., $\bar{X}_{1j} - \bar{X}_{2j} = c$, are by chance. A statistical test like the Student's t-test (Kanji 2006, p. 8) can show if, despite the differences in the means of the observed distribution, the samples for the two models can still be samples of the same underlying distribution (identical means). The Student's t-test requires the underlying distributions of both distributions to be normally distributed. As this is not guaranteed here, a more generic non-parametric test that does not require any assumptions to hold for the underlying distribution of the samples is preferable, such as the two sample Kolmogorov-Smirnov test (KS-test; Massey 1951). It relies on the empirical

cumulative distribution functions (stepfunction S_n) which for a random variable X with n observations is defined as (Massey 1951):

$$S_n(x) = k/n \quad (6.1)$$

where k is the number of observations less or equal to x . From the S_n calculated for each sample, a test statistic is then calculated as:

$$D = \sup |S_{n,1}(x) - S_{n,2}(x)|. \quad (6.2)$$

Using the distribution of the test statistic D whose CDF is approximated as:

$$F_{KS}(x) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 x^2}, \quad (6.3)$$

and given a significance level α (e.g., 5%), the null hypothesis is rejected if $p_{KS} = 1 - F_{KS}(D^*) < \alpha/2$ (half because a two sided test is used), where for continuous X (Stephens 1974):

$$D^* = (e_n + 0.12 + 0.11/e_n)D \quad (6.4)$$

with

$$e_n = \sqrt{n_1 n_2 / (n_1 + n_2)} \quad (6.5)$$

Statistical significance of the model pairs

Running the KS test on the loss distributions at 1, 2, 5, and 10% PoE, all model combinations yield test statistics KS and p_{KS} . With a significance level of $\alpha = 0.05$, almost all differences are significant. Table 6.3 lists only non significant differences. For the full table, please refer to Table F.1 in the Appendix.

As can be seen from Table 6.3, the difference between a Poissonian model with immediate repair and one with delayed repair is not significant. For the *ETAS* model, the differences between models with *conservative* and *optimistic* delayed repair are found to be non-significant. The last non significant difference is identified for a Poissonian model with only a *conservative* delayed repair model and one with additional state-dependency, but only at a PoE of 5%, with the test being close to the chosen significance-level ($p_{KS}=0.05$). Another boundary case is the difference between the *ETAS* model with the *conservative* repair and one with the *optimistic* repair at 10 % PoE, whose distributions are shown in Figure 6.9. Indicating that the difference between the optimistic and conservative model is close to significance.

A simple tool to compare two distribution samples is a quantile-quantile (qq) plot, for which one computes quantiles of each distribution and plots them against each other. Figure 6.10, shows such a qq-plot for the two distributions in Figure 6.9. The closer the quantiles, the closer they are to the red 45° line in the figure. As one can see, the quantile ratios are very close to this line, however, at both ends, low and high, some scatter can be observed, which however is not significant according to the test.

Table 6.3.: Pairwise model combinations whose differences in the loss distribution at different levels of exceedance probability (PoE[%]) are non-significant according to a two-sided Kolmogorov-Smirnov test (KS , p_{KS}) with a significance level $\alpha = 0.05$; d is the size of effect (Cohen 1988) for the mean values.

Model high median	Model low median	PoE	KS	p_{KS}	d
50y DR(1,0.5)+SD	50y FR(1)	1	0.04	0.53	0.01
50y DR(1,0.5)+SD	50y FR(1)	2	0.02	0.99	0.01
50y FR(1)	50y DR(1,0.5)+SD	5	0.03	0.79	0.01
50y FR(1)	50y DR(1,0.5)+SD	10	0.03	0.75	0.01
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(-0.1,0.4)+SD	1	0.05	0.21	0.00
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)+SD	2	0.04	0.46	0.01
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)+SD	5	0.04	0.28	0.03
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(-0.1,0.4)+SD	10	0.06	0.05	0.09
50y DR(1,0.5)	50y DR(1,0.5)+SD	5	0.06	0.05	0.02

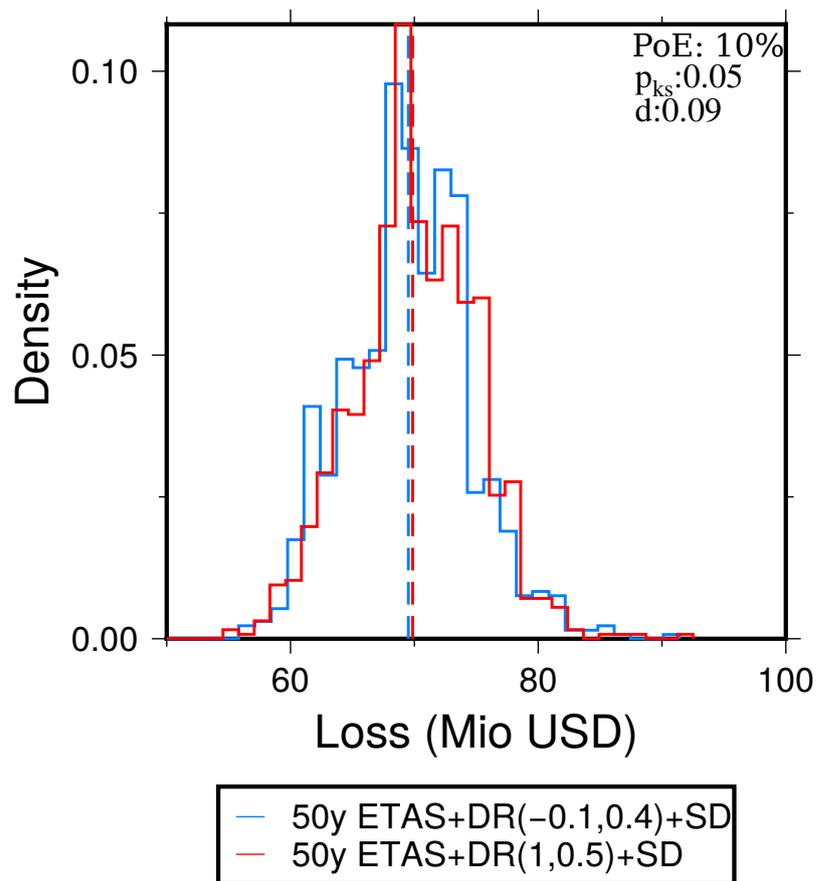


Figure 6.9.: Boundary case identified for the difference in loss distributions for 10% PoE between the $ETAS+DR(1,0.5)+SD$ and $ETAS+DR(-0.1,0.4)+SD$ model. With an $\alpha = 0.05$, this difference is considered non-significant using a two sided Kolmogorov-Smirnov test ($p_{KS} = 0.05$). Cohen's d describes the size of effect, i.e., the difference of the sample means normalized by the standard deviation of the pooled samples.

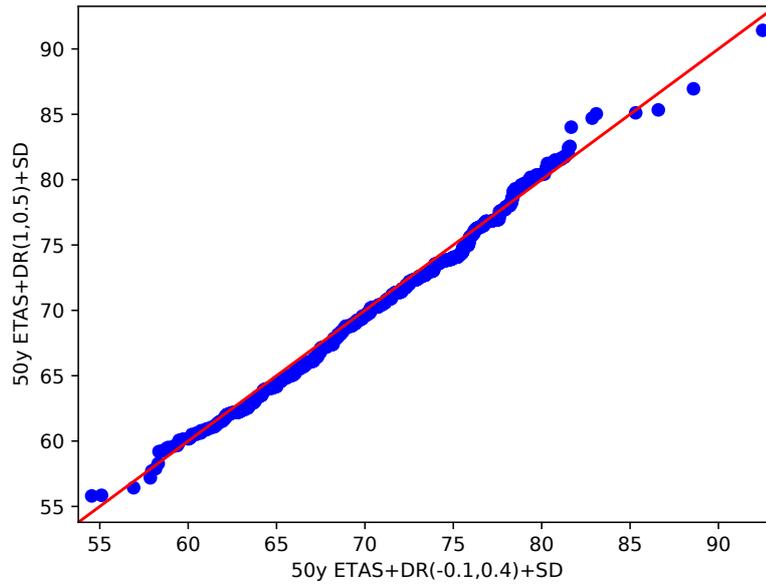


Figure 6.10.: Quantile vs quantile plot of the two loss distributions at 10% PoE for the $ETAS+DR(1,0.5)+SD$ and $ETAS+DR(-0.1,0.4)+SD$ model (Figure 6.9), in which case the difference between both is close to significance with $p_{KS} = 5\%$ ($\alpha=5\%$). Red line corresponds to identical quantiles.

Size of effect for different modeling decisions

The KS test only determines if a difference in the sample distributions, as small as it may be, is statistically significant or not. A more practical measure for the relevance of this difference is the size of effect. It can be calculated as Cohen's d (Cohen 1988) with a pooled standard deviation of the two sample distributions:

$$d(X_1, X_2) = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(s(X_1)^2 + s(X_2)^2)/2}} \quad (6.6)$$

where \bar{X}_i and $s(X_i)$ are mean and standard deviation of X_i . and can be interpreted as a value of non-overlap of distributions, e.g., by 0% ($d=0$), 55.4% ($d=1.0$), 81.1% ($d=2.0$).

The largest average Cohen's d is found between the *conservative* ETAS model and the classical model, ranging from 3.09 at 1% until 5.88 at 10%. Ranging between 3.09 and 5.89, the size of effect for the full model and a Poissonian model with delayed repair and state dependency ($DR(1,0.5)+SD$) ranks second. The third in terms of size of effect is the mean difference between the model with *optimistic* repair and the classical model, with d ranging from 3.01 at 1% PoE until 5.74 at 10% PoE. The lowest, but statistically significant, size of effect ranges from 0.13 until 0.21 (varying per PoE) for the differences in the distributions of the *optimistic* ETAS model and the *pessimistic* ETAS model. Varying between 0.03 and 0.27, the second lowest rank for the size of effect is observed for the difference between the Poissonian models with fixed repair $FR(1)$ and with delayed repair without state

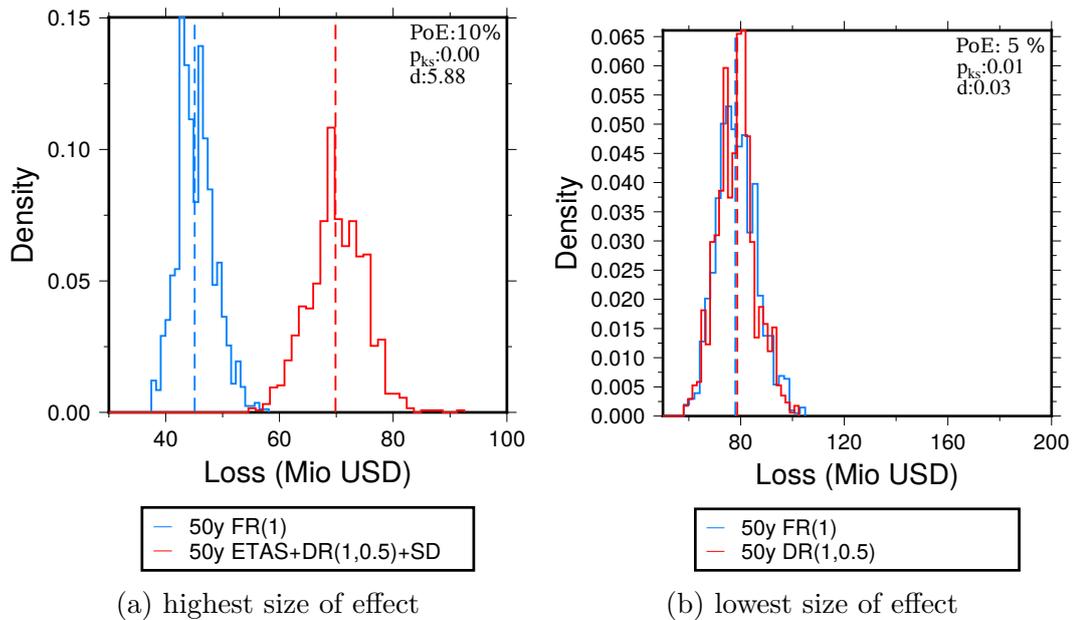


Figure 6.11.: Loss distributions for the two model pairs which yielded the highest and lowest significant size of effect among all model pairs. The plot on the left shows the highest observed size of effect with $d = 5.88$; on the right with $d = 0.03$ the lowest significant size of effect; p_{KS} : p-value for the Kolmogorov-Smirnov test.

dependency $DR(1,0.5)$. The third ranking lowest size of effect is observed for the difference between the *conservative* and the *pessimistic* ETAS model. Figure 6.11 shows the distributions for the highest and lowest size of effects, and corresponding test values, as well as size of effect. Meanwhile, Figure 6.12 presents two qq-plots for the same model combinations, showing almost no deviation from the red line in the case of the lowest observed size of effect and a very large deviation for the case of the largest observed size of effect.

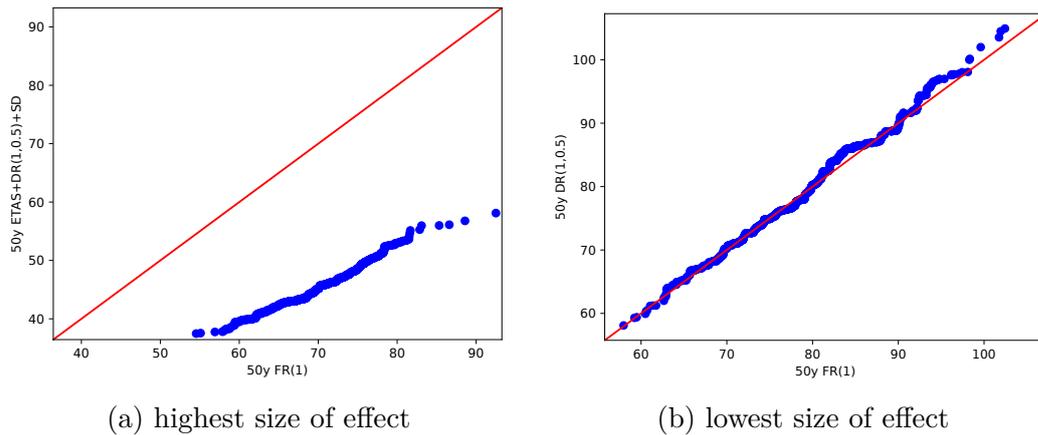


Figure 6.12.: Quantile vs quantile plots for the two model pairs which yielded the highest and lowest significant size of effect (see Figure 6.11).

6.4. Single year observation period

While 50 year observation periods are suitable to highlight the long term effects of the considered model components, such a long period is rarely used for risk assessments in practice. Especially in the insurance industry, annual loss models are more common, as this aligns with the fiscal year and many reinsurance contracts have a renewal period of about 1 year. Due to the shorter period of observation, usually also lower probabilities are considered than for 50 years. A common lower end threshold is 0.01% exceedance probability in a year, i.e., on average one event in 10'000 years. In order to resolve this probability for an EP-curve, at least 10'000 simulations are required. Ideally, to obtain results with similar reliability as the 50 year simulations, at least 500 simulated years should support the lowest level, which sets the number of simulations to 5 million. If the uncertainty is to be determined again through a 10% bootstrap, this adds another factor of 10. However, considering the purpose of the presented application to demonstrate the frame work with the goal of identifying tendencies, 5 million simulated years are beyond the scope here.

Differences in total losses

Thus, accepting higher uncertainties in the absolute numbers, Figure 6.13 shows annual EP-curves calculated without bootstrapping from 25'000 one year simulations. On the left the curve for a model with full sequences (*ETAS*), the *conservative* repair model ($DR(1,0.5)$), and state-dependent fragility (*SD*) is shown. As the comparison with a model under classical assumptions ($1y FR(1)$) shown next to it indicates, the tendency replicates the previous observations for 50 years. Over the full period range, an increase of simulated losses for the full model is found, when compared to the classical one. Please note that due to the low number of simulations, under-sampling effects are very pronounced for PoEs with less than 500 years support, i.e., below 2% PoE (indicated by the marked area). This can be seen from the kinks in the curves and also in the apparent saturation of the classical loss at 340 Mio USD

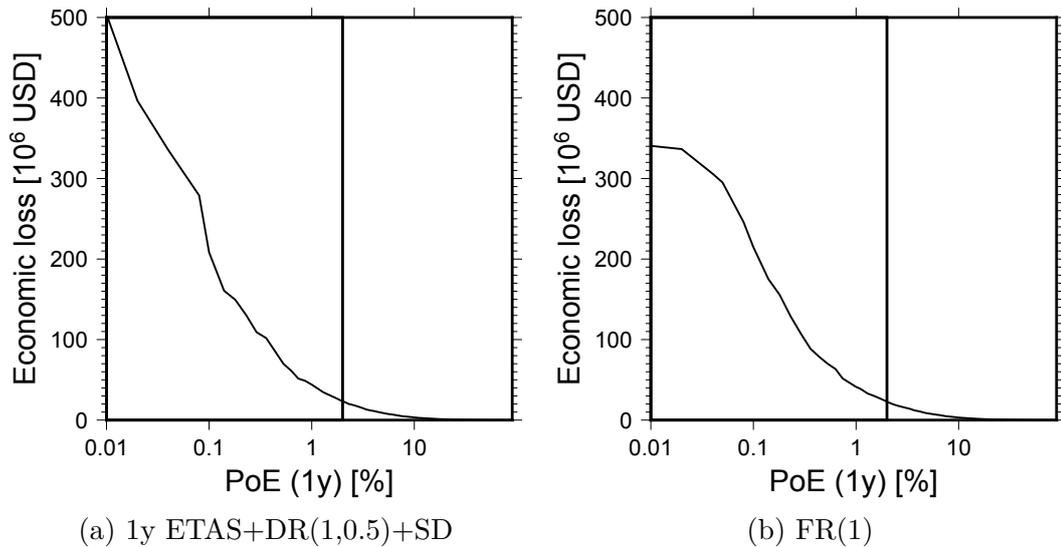


Figure 6.13.: Annual loss exceedance probability (PoE) curves for 1'000 simulated buildings in Nablus, without bootstrapping, calculated from 25'000 simulations for a single year of seismicity. On the left curves from a time- and state-dependent model $1y\ ETAS+DR(1,0.5)+SD$ and on the right side one for classical assumptions $1y\ ETAS+DR(1,0.5)+FR(1)$. Each curve is based on 25'000 years of simulated seismicity. Hence, the lower probability end corresponds to only 2.5 simulated years and very uncertain. The gray marked areas indicate the parts of the curves that are supported by less than 500 simulated years.

below 0.1% PoE . Furthermore, simulating only a single year, the *ETAS* model is strongly conditioned by the input catalog.

Figure 6.14 shows the residuals between both models. While for high PoEs only slight differences occur, as most years in this range don't cause aftershocks above $Mw = 4$, a clear and rapidly increasing difference can be seen below 1 % PoE. The gray marked area highlights the part with low support (< 500 years) which can also be seen from strong fluctuations in the residuals. Between 0.2 and 0.1 % PoE the full model seems to yield lower losses than the classical model. Due to the low support, no reliable general conclusions can be drawn, but the reasons observing this drop here can be explained as it aligns with the range where multiple Poissonian, potentially damaging, events occur within a year. In case of delayed repair, already damaged buildings are not damaged again in the model with delayed repair, and with only moderate damage, the state dependency of the fragility is still not very pronounced. Furthermore, the immediate repair causes multiple additional low level transitions in the model with immediate repair as previously observed also for 50 year periods.

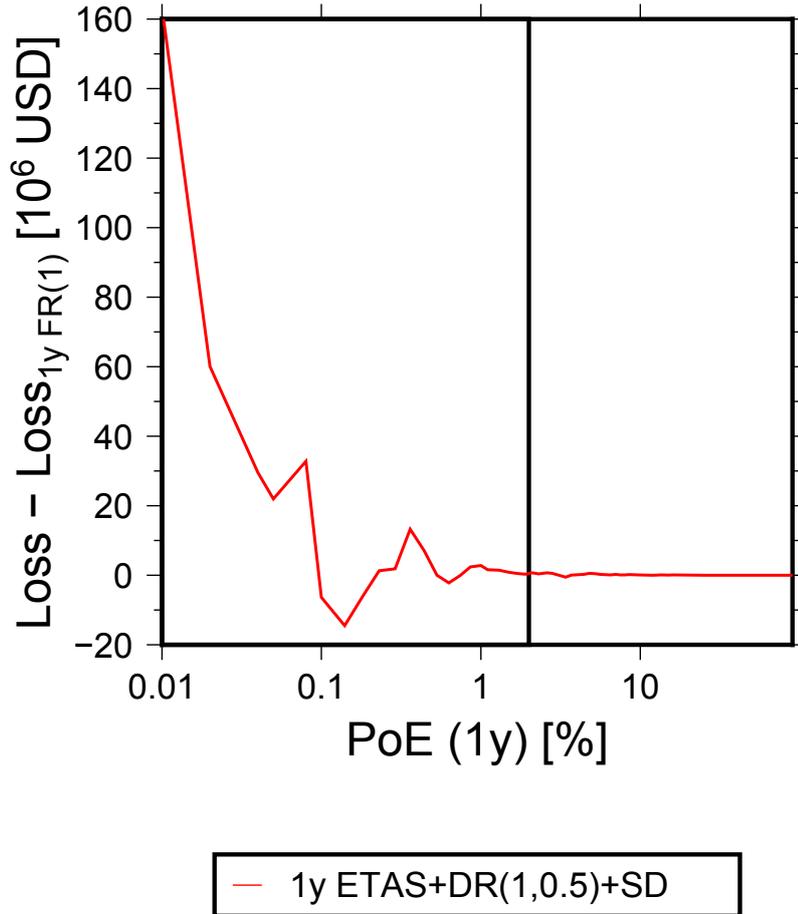


Figure 6.14.: Residual between the annual loss exceedance probability (PoE) curves for a $1y$ $ETAS+DR(1,0.5)+SD$ model and a classical $1y$ $FR(1)$ model, both EP-curves calculated without bootstrapping from 25'000 simulations for a single year of seismicity, the lower probability end thus corresponds to only 2.5 simulated years and is very uncertain. The gray marked areas indicate parts of the curves that are supported by less than 500 simulated years.

Table 6.4.: Simulation runs for a 1 year observation period under different modelling conditions employing the developed frame work.

ID	$T_{obs}[a]$ ¹	ETAS	Repair ²	State dependency
1y FR(1)	1	✗	FR(1)	✗
1y real MS+FR(1)	1	✓(MS) ³	FR(1)	✗
1y ETAS+DR(1,0.5)+SD	1	✓	DR(1,0.5)	✓
1y ETAS+FR(0)+SD	1	✓	FR(0)	✓
1y ETAS+FR(0.00365)+SD	1	✓	FR(0.0000623)	✓

¹ Observation period, i.e., simulated period after input catalog.

² DR: time delayed repair with underlying normal distribution of lognormal model (μ , σ); FR: fixed probability repair (probability=1).

³ MS: ETAS model was used and mainshocks extracted.

Potential for simplifications

Considering annual loss simulations, more questions regarding the model decisions arise, the first being when only single years are simulated, is a complex time delayed repair model required or can one just assume a very small fixed repair probability (Cesare et al. 1992), or even assume no repairs to happen at all (e.g., Tesfamariam and Goda 2017; Papadopoulos and Bazzurro 2018). A second question that is briefly investigated below regards the previously made statement that using Poissonian events, but not the actual main shocks of the sequences, is approximately equivalent to a main shock model (see section 5.1.4). If they are not approximately equivalent, a systematic difference might bias the results, and thus the comparison between seismicity from the Poissonian and the ETAS models, especially if pseudo random numbers drive the calculations.

A set of additional single year simulations, listed in Table 6.4, were run in order to answer these questions. Figure 6.15 shows EP-curves for the different models relative to the EP-Curve of the classical model $1y FR(1)$. The first observation when these results are compared to those of the 50 years simulation are strongly pronounced variations. This is due to the lower support, where parts of the curves which are supported by less than 500 years of seismicity are marked by a gray area again.

As previously, red curves indicate *ETAS* models, but in contrast, the only blue curve ($1y real MS+FR(1)$) represents an ETAS model which was reduced to the main shocks (MS) of the sequences. For the latter model, the average trend over all PoEs indicates an increased loss of a few percent when compared to the Poissonian model ($FR(1)$). This translates to a slight overestimate of the difference between Poissonian and ETAS seismicity models in the previous section, although, due to the longer periods of 50 years used therein, the differences are likely to be less pronounced. The only exception to the on average small bias observed when considering real mainshocks is a strong peak between 1 and 0.2% PoE where the main shock reaches 30% higher losses, than the purely Poissonian trigger model. It agrees with a peak that occurs in all *ETAS* models, which indicates that the simulations that contribute to this PoE range in the EP-curves show many strong triggered events, whose magnitudes exceed that of the Poissonian triggering events (used for the $FR(1)$ model). The combination of immediate repair and several mainshocks in the real mainshock model causes a strong loss amplification. Below about 0.15% PoE, this behavior is not observed, reducing the difference to the order of few percent again.

No repair or constant daily repair probability model

In the case of the models with delayed repair, both, time-dependent DR and constant $FR(p \neq 1)$ probability models, dampen the strong increase observed for the real mainshock model at 1% PoE. Despite strong fluctuations, a positive trend towards higher losses is observed for all *ETAS* simulations. Below 0.1% PoE, this

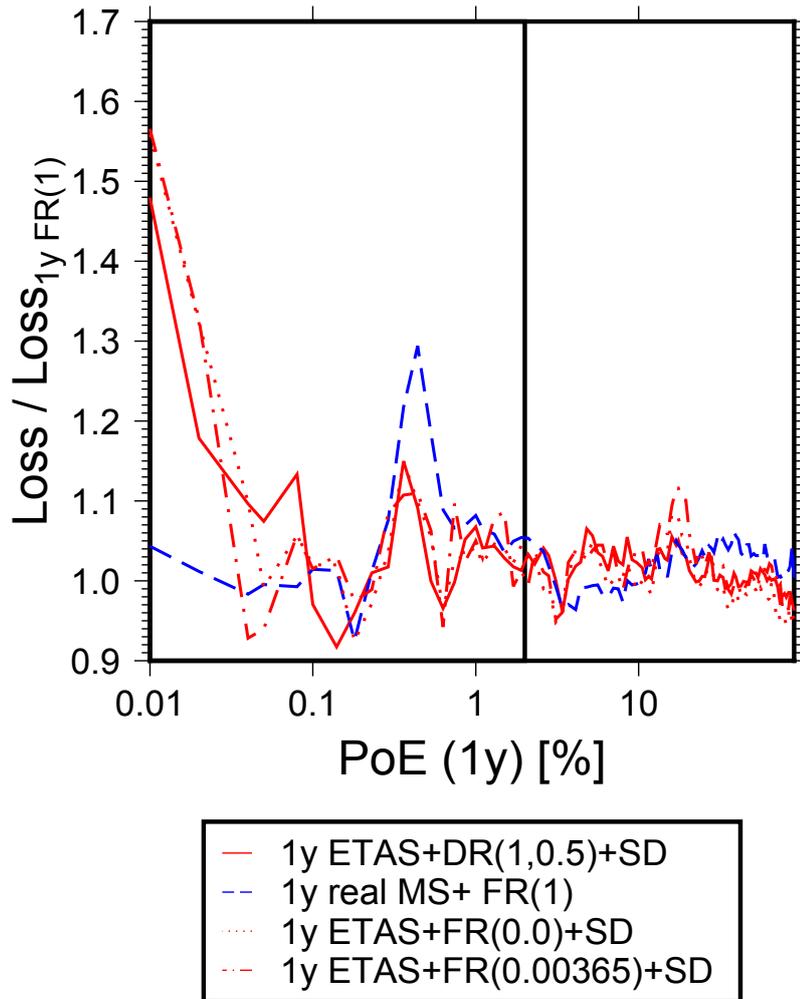


Figure 6.15.: Annual loss exceedance probability curves for models in Table 6.4, considering full seismic sequences *ETAS*, real mainshocks extracted from the *real MS*, delayed repair of lognormal shape $DR(\ln(\mu) = 1, \ln(\sigma) = 0.5)$ [years], and fixed repair with daily probability p ($FR(p)$), all relative to a classical Poissonian annual risk model ($FR(1)$); relative loss equal to 1). Each curve is based on 25'000 years of simulated seismicity, the lower probability end thus corresponds to only 2.5 years and is very uncertain. The gray marked areas indicate parts of the curves that are supported by less than 500 simulated years.

strongly increases, reaching ratios that indicate up to 56% higher losses with respect to the classical model for a model without any repair ($ETAS+FR(0)$). Until about 0.2% PoE, the full model ($ETAS+DR(1,0.5)+SD$), deviates only slightly from those with the simpler assumptions without repair and a model that employs a fixed daily repair probability of $p = 0.0000623$ ($ETAS+FR(0.0000623)$), which is calculated evaluating the *conservative* repair model ($DR(1,0.5)$) daily over one year and averaging resulting (incremental) probabilities. Please note also, that part of the variations can be attributed to the low support ranging between 50 and 250 simulated years in this part of the curves.

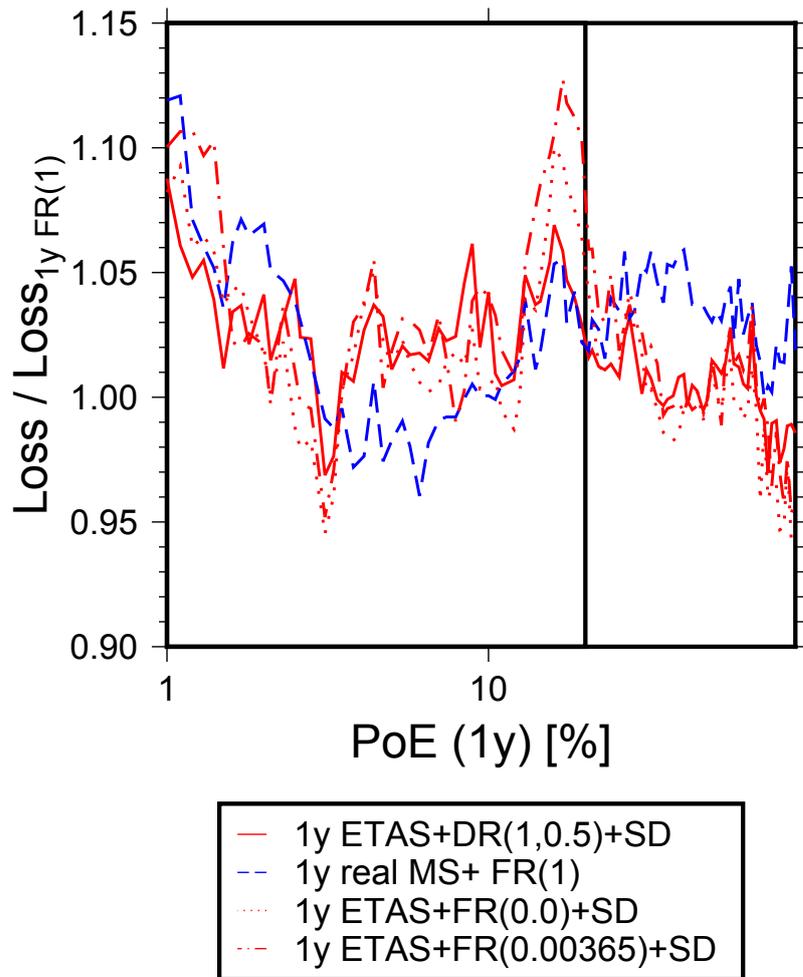


Figure 6.16.: Annual loss exceedance probability curves as in Figure 6.15, but determined using a 10% bootstrap, hence the probability range is limited at 1%. The gray marked areas indicate parts of the curves that are supported by less than 500 simulated years.

Statistical significance

To determine the statistical significance of the observed differences as previously described in section 6.3, the loss distributions, and thus bootstrapping are required. As previously seen with the 50 years observation period, generating 10% bootstrap samples of all simulations (2'500 simulations each) yields EP-curves defining loss distributions for the considered PoE levels. Due to the low number of simulations, however, the statistical analysis is restricted to 5, 10, and 20% PoE, where only the latter is supported by 500 years, the others by 125 and 250 simulated years, respectively. The reliability is therefore below that for the 50 year observations. Figure 6.16 shows the median EP-curves resulting from this bootstrapping. Please note the changes in scales with respect to Figure 6.15. The figure has been limited to a minimum PoE of 1%, corresponding to 2.5 years of supporting simulations, the gray area once more indicating PoEs supported by less than 500 years. Applying the statistical Kolmogorov-Smirnov test as previously described and using a significance level of 5%, indicates significance for most pairwise model combinations. Table 6.5

Table 6.5.: Pairwise model combinations with 1 year observation period whose differences in the loss distribution at different levels of exceedance probability (PoE[%]) are non-significant according to a two-sided Kolmogorov-Smirnov test (KS, p_{KS}) with significance level $\alpha = 0.05$; d is the size of effect (Cohen 1988) for the mean values.

Model high median	Model low median	PoE	KS	p_{KS}	d
1y ETAS+FR(0.0)+SD	1y FR(1)	10	0.04	0.31	0.03
1y ETAS+FR(0.0000623)+SD	1y ETAS+FR(0.0)+SD	5	0.03	0.68	0.05
1y ETAS+FR(0.0)+SD	1y real MS+ FR(1)	10	0.05	0.13	0.06
1y ETAS+DR(1,0.5)+SD	1y real MS+ FR(1)	20	0.06	0.07	0.02
1y FR(1)	1y real MS+ FR(1)	5	0.04	0.31	0.07
1y real MS+ FR(1)	1y FR(1)	10	0.06	0.07	0.03
1y ETAS+FR(0.0)+SD	1y ETAS+DR(1,0.5)+SD	5	0.04	0.31	0.02
1y ETAS+FR(0.0000623)+SD	1y ETAS+DR(1,0.5)+SD	5	0.03	0.79	0.02
1y ETAS+DR(1,0.5)+SD	1y ETAS+FR(0.0000623)+SD	10	0.05	0.09	0.00

list the model combinations for which no significant difference was found (a complete table of test results is presented in Appendix F.2). The previously observed differences between the real main shock model (*1y real MS+FR(1)*) and the model with Poissonian seismicity (*1y FR(1)*) are found to be non-significant with the exception of the 20% PoE level, and a tendency for increasing significance towards higher PoEs, which is a result of the higher support and raises doubts about the non-significance at lower PoEs. According to the test, the loss distributions at 5 and 10% PoE of the model with *conservative* time-dependent delayed repair and the constant low daily probability model are not significant. This agrees with the previous observations of Figure 6.15, where a difference was only developing at the low PoE range. At 5% PoE, the model without any repair is not significantly different from the model, with the constant low probability. The former's loss distribution was also not different from the classical model with immediate repair *FR(1)* at 5% PoE. This falls in the PoE range, where a drop of all *ETAS* models is observed after the peak in Figure 6.16 and where the no-repair model yields the lowest losses among the *ETAS* models, as subsequent events cause no further damage, whereas in the other models, some buildings are repaired and damaged again.

Size of effect

Generally, the sizes of effect (Cohen's d) for the annual simulations are lower than those observed for the 50 year simulations. Figure 6.17 shows the highest and lowest size of effect observed for all considered PoEs and model combinations. The left shows with $d = 0.54$ the highest observed value for the model with constant low daily repair probabilities with respect to the classical model at a PoE of 20%. On

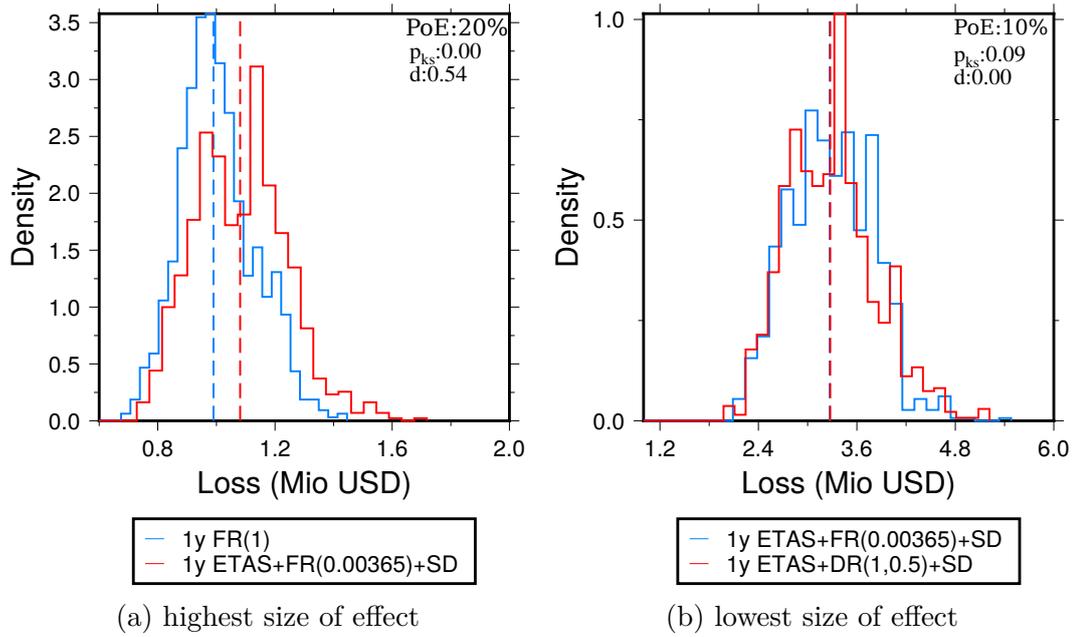


Figure 6.17.: Loss distributions for the two model pairs which yielded the highest and lowest size of effect among all model pairs. The plot on the left shows the highest observed size of effect with $d = 0.54$; on the right with $d = 0.00$ the lowest (non-significant) size of effect; p_{KS} : p-value for the Kolmogorov-Smirnov test.

the right side of Figure 6.17, the model with a low constant repair probability is shown again, this time showing the lowest size of effect ($d = 0$) for a PoE of 10% when compared to the model with the *conservative* time-dependent repair model.

On average, among the model pairs where differences were significant for all considered levels of PoE, the largest effects on the mean losses are observed for the fixed low probability repair model and the classical Poissonian model with d ranging between 0.17 and 0.54. With a size of effect from 0.21 until 0.33, the same model's difference with respect to the real mainshock model ranks second. Considering the time-dependent repair model, yields an effect size between 0.16 and 0.2 when compared to a classical estimate and thus ranks third. The on average lowest size of effect on the mean is observed between the model with time-dependent repair and a model without any repair, ranging between 0.02 and 0.14. The second-lowest effect, with d from 0.03 until 0.22, results from considering real main shocks with respect to a classical Poissonian model. The difference between the fixed low daily repair probability and the time-dependent repair model ranks third, with d spanning from 0 until 0.35 between. Note that in addition, the differences for the latter two were found insignificant for most tested cases.

7. Validation of results

Due to the many simplifications employed, the reliability of the risk estimates shown in the previous chapter is relatively limited. While the validation of the used seismic hazard models is feasible to some extent (e.g., Mak et al. 2014), seismic risk models are difficult to validate, due to the very dynamic nature of exposure and to some extent the changing vulnerability of structures. Frequently, insurance companies are able to assess the validity of their risk models to some extent as they have records of liabilities and a fixed exposure defined by the insured values. Lacking both for the Dead Sea region, a full validation of the presented model is thus unfortunately not possible. Although the model is only intended for demonstration purposes, this chapter contains several checks of individual components to provide an indicator for potential biases in the results. None of the tests is intended to be exhaustive, but should only suffice as a first-order check.

7.1. Validity checks for the seismicity

In order to determine the validity of the seismicity model employed in the simulation framework, a first check regards the Gutenberg Richter law for the seismicity generated by all 10'000 ETAS simulations. Up to a rounding error of the order of 10^{-5} , the input parameters $b = 1.12$ and daily rate $\mu = 0.0053$ is obtained from the stochastic catalogs. To investigate the variability of the simulation results, the bootstrap method (Efron 1979) was employed, choosing 10% random samples from the simulations (which are all independent realizations of the same random process). Using these randomly chosen 1'000 simulations of 50 years of seismicity and spanning a grid with 0.1 by 0.1 degrees spanned over the region, earthquake counts and released seismic moment are calculated for each of the bootstrap samplings. Figure 7.1 shows the mean and standard deviation of the earthquake count per 0.1 by 0.1 degree grid cell for all bootstrap samples. Comparing these to the input background probability (Figure 5.2), one observes that it replicates the given distribution. The circular smoothing effect comes from the isotropic placement of aftershocks as only Poissonian events are placed according to the background probability and the aftershocks follow the linearized density model (equations 4.7 -4.10).

On the same grid, the seismic moment of all generated earthquakes is estimated (Hanks and Kanamori 1979):

$$\log_{10}M_0 = 1.5M_w + 16.1 \text{ [dyn cm]}. \quad (7.1)$$

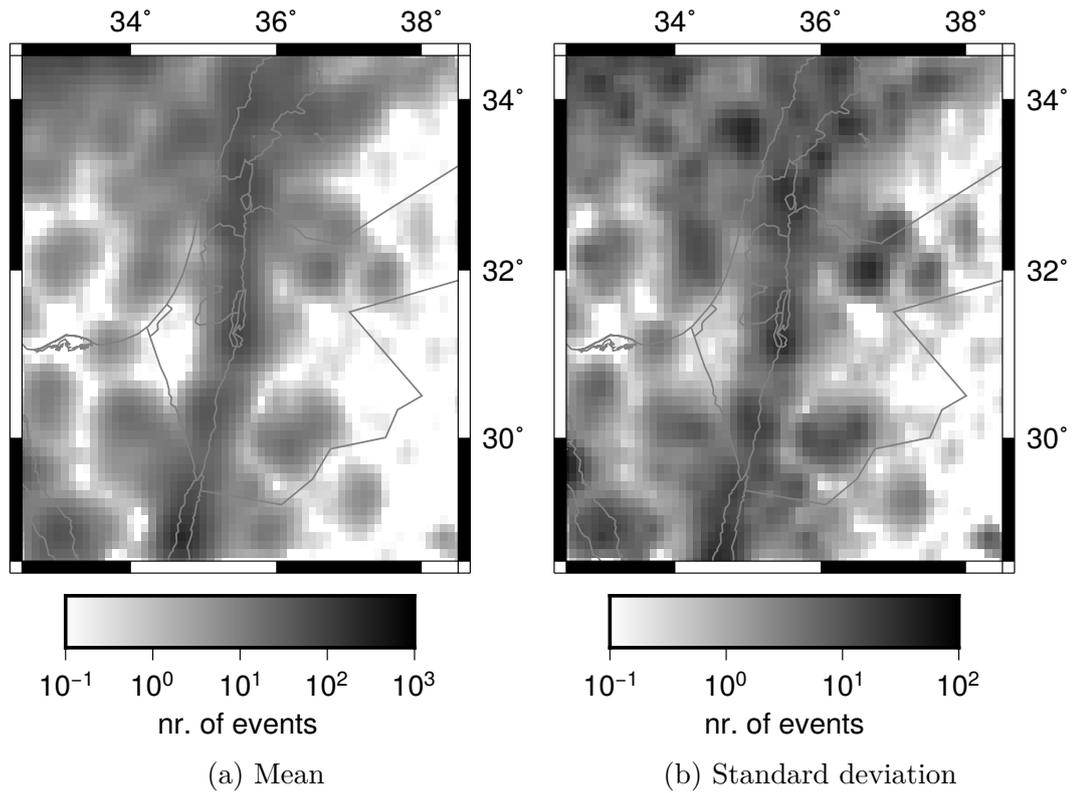


Figure 7.1.: Mean and standard deviation of earthquake counts over a 50 years period on a regular 0.1 by 0.1 degree grid determined using a 10% bootstrap sampling approach.

Again, using the bootstrap method, the resulting mean and standard deviations in $[PNm]$ for each grid cell are shown in Figure 7.2.

Figure 7.2 underlines the previous observation, and in addition indicates by darker colors those areas where larger events were generated in the simulation. The standard deviation would indicate if these places show consistently larger events, i.e., indicating large variations at locations with low means or vice-versa. As the standard deviation replicates the spatial pattern of the means this is not the case and is not surprising, as a spatially homogeneous Gutenberg-Richter law was employed.

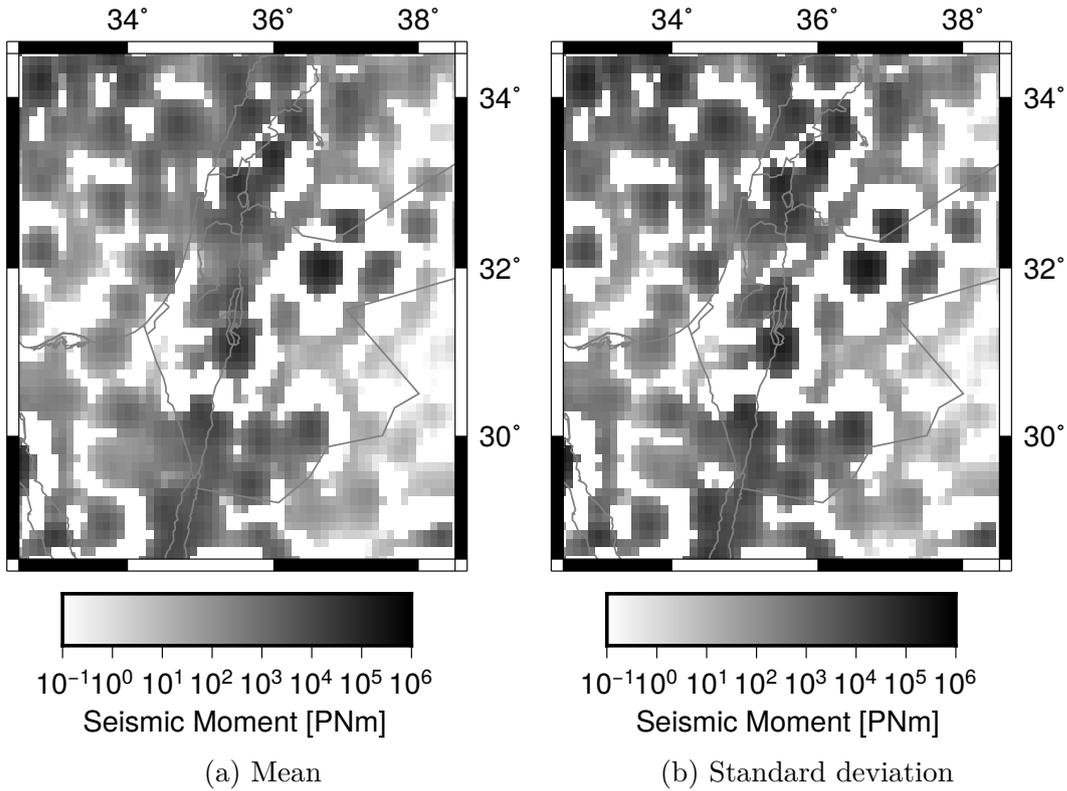


Figure 7.2.: Mean and standard deviation of the seismic moment over a 50 years period on a regular 0.1 by 0.1 degree grid determined using a 10% bootstrap sampling approach.

7.2. Hazard

For the hazard, only Poissonian hazard estimates exist for Nablus. A recently published model is the Earthquake Model Middle East (EMME; Şeşetyan et al. 2018). Figure 7.3 shows the 5, 50, and 95 % quantiles hazard curves from this model on rock ($v_{s30} = 800m/s$) in comparison to a hazard estimate from Poissonian seismicity derived from using the simulation framework and seismicity model as previously shown in Figure 5.6. The estimates of the EMME model suggest a higher median hazard, for example at 0.001 PoE the difference in the median is about 20%. The uncertainty at low PoEs indicated for the EMME model is lower than those obtained here for the simulation results.

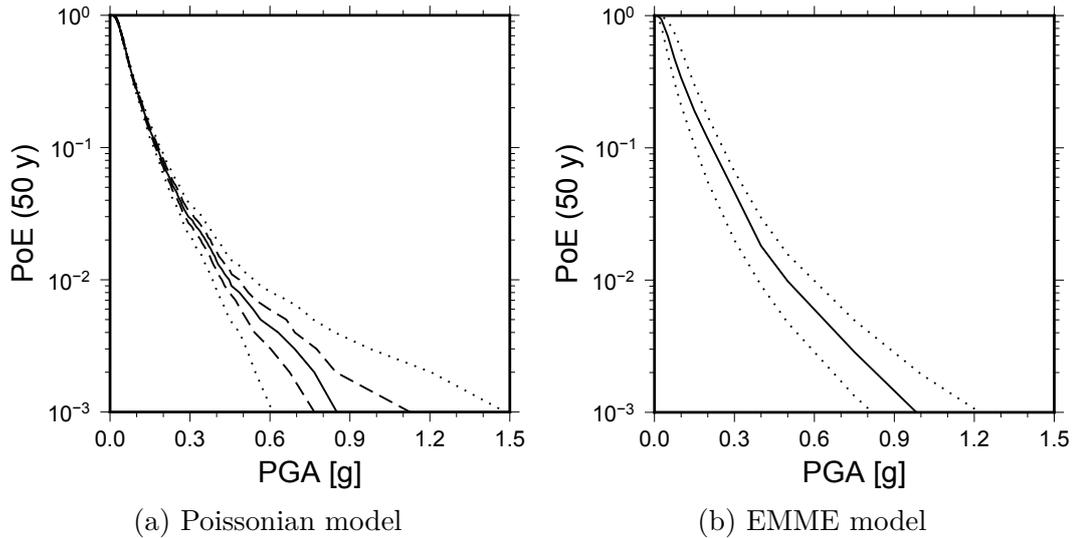


Figure 7.3.: On the left: Poissonian median hazard curve (solid) and 5, 25, 75, and 95 % quantiles for a single site in central Nablus (35.268 E, 32.220 N, $v_{s30} = 760m/s$) as obtained with the simulation framework. The right side shows Poissonian median and 5 and 95% quantile hazard estimates for Nablus from the EMME model (Şeşetyan et al. (2018); available from www.efehr.org) on rock ($v_{s30} = 800m/s$).

7.3. Exposure model validation

To the best knowledge of the author, the only other publicly available exposure model for Nablus is based on a study by Grigoratos et al. (2016), who performed a small exposure and fragility study on 112 buildings in Nablus. Regarding exposure, they found 96% cast in place RC frame buildings without earthquake resistant design and 4% masonry low rise structures. Although according to Grigoratos et al. (2016) local engineers estimate that about 5% of the building stock should be recently built RC dual system buildings with seismic design, their data set did not contain any of these types. The results of the exposure analysis presented within this thesis generally agree with the results from Grigoratos et al. (2016). Indeed, there are 23 buildings in the data set collected here which have been assigned as being ductile and one even as base isolated, corresponding to a total of 2.9%. This is slightly below the 5% estimated by the local experts.

7.4. Bias in fragility models

The developed state-dependent fragility models employed in the time and state-dependent simulations are not realistic. A short comparison of the methodology when applied to an experimental building and other approaches is given in Appendix C.5. Specifically, regarding the fragilities determined for Nablus, only limited comparisons are possible as only Grigoratos et al. (2016) provides some local fragility estimates. These were only for intact buildings and were, unfortunately for

the comparison, in EMS-98 damage grades instead of HAZUS ones. A drawback for

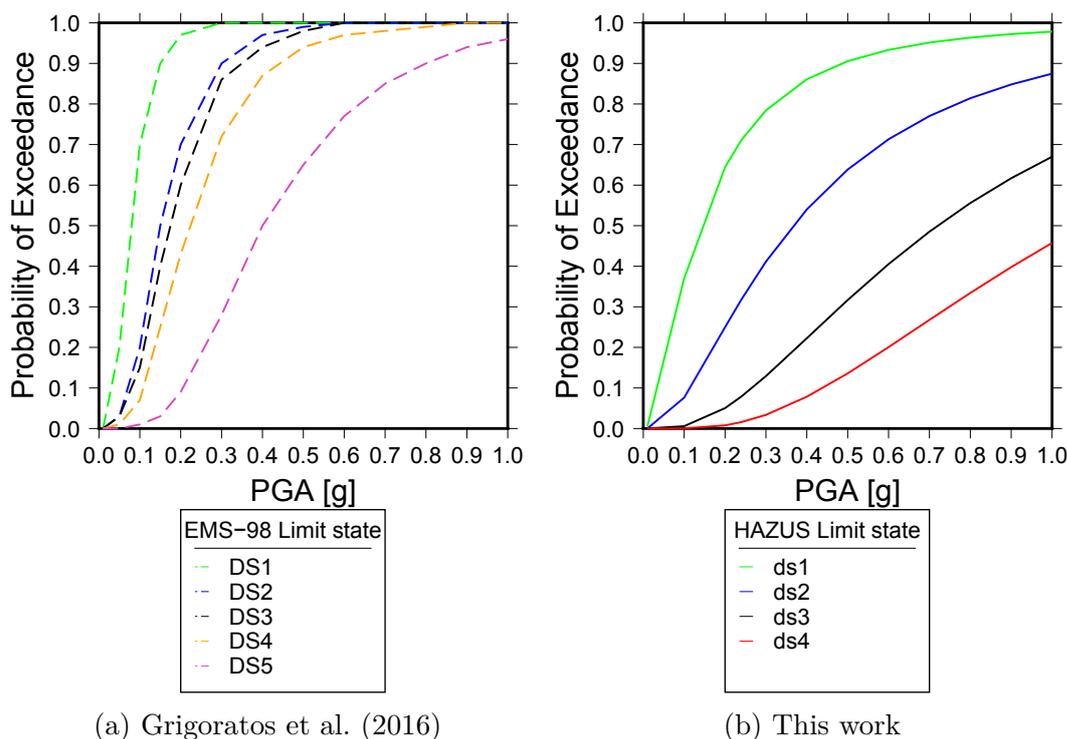


Figure 7.4.: Fragility model for an intact 7 floor RC Frame building in Nablus (left) as proposed by Grigoratos et al. (2016) using EMS-98 damage states (DS_i), and (right) as determined in this study using HAZUS damage states (ds_i; averaged over all 7 floor models).

the comparison is that the EMS-98's damage grades are five instead of four as in HAZUS. The difference results from the fact that EMS-98 distinguishes between partial collapse (*Very heavy*), i.e., collapse of non-load bearing elements, and complete collapse (*Destruction*) for what is uniformly termed *complete* damage in HAZUS. Therefore, only the first three damage grades can be directly compared as these are equivalent (Milutinovic and Trendafiloski 2003, p. 17). However, Figure 7.4 shows a strong underestimation of the fragilities used for this study when compared to the model of Grigoratos et al. (2016), independent of the damage state. This is underlined when comparing the exceedance probabilities provided by Grigoratos et al. (2016) for 2, 4, 7 and 10 storey buildings at a ground motion level of $PGA = 0.24g$ with the averaged aggregates of the models used here (Table 7.1).

Unfortunately, the damage state dependency effect can not be validated directly as such data is not available for Nablus nor for the experimental ICONS buildings. However, comparing the modeled probability increases for Nablus between 1 and 26% to results, e.g., obtained in Trevlopoulos and Guéguen (2016), which find between 5 and 30% probability increases for low-code, low- and mid-rise moment-frame structures, the estimates seem reasonable. Nevertheless, due to the employed linear models an underestimate of the effect can not be ruled out.

Table 7.1.: Average exceedance probabilities of HAZUS damage states ds_1 , ds_2 , ds_3 , and ds_4 for building types aggregated by number of stories. For comparison, the exceedance probabilities as obtained by Grigoratos et al. (2016) for buildings in Nablus using EMS-98 damage states (DS1-5) are provided in brackets.

building type	ds_1 (DS1) [%]	ds_2 (DS2) [%]	ds_3 (DS3) [%]	ds_4 (DS4, DS5) [%]
2 f	67 (89)	21 (52)	3 (46)	1 (35, 8)
4 f	75 (97)	32 (75)	12 (69)	4 (55, 15)
7 f	71 (99)	32 (80)	8 (74)	2 (57, 16)
10 f	77 (98)	38 (77)	10 (70)	2 (52, 12)

7.5. Fully empirical Poissonian risk model for Nablus

As a further attempt to partially validate the results of the simulations, a seismic risk model relying on purely empirical components is developed for Nablus. Such a model can be based on macroseismic intensity (EMS-98) which has several benefits in the case of Nablus. It allows the hazard to be better constrained, as pre-instrumental observations related to ground motions are more readily available for the region than instrumental records. The hazard estimates from such an empirical model can thus be considered statistically more stable. Furthermore, the EMS-98 defines vulnerability classes, which are based on large scale statistics that describe the expected damage for buildings with varying seismic resistance. The original formulation of EMS-98 qualitatively describes the number of observed damage levels using terms like *few*, *many* and *most*. These qualitative descriptions were turned into a quantitative vulnerability index measure by Giovinazzi and Lagomarsino (2004). Determining the vulnerability class distribution for Nablus and following Giovinazzi and Lagomarsino (2004) allows one to build an empirical risk model. This model can explicitly consider related uncertainties and will be implemented in the OpenQuake software (Pagani et al. 2014). Unfortunately, such a EMS-98 based model is not able to produce comparable time- and state-dependent risk estimates. However, it may yield a baseline for a classical risk estimate. This baseline can be used to evaluate, at least partially, how well risk estimates obtained from the proposed simulation framework, where time and state-dependency are not considered ($FR(1)$ model in Table 6.1), correspond to this baseline.

While for the simulation framework only aleatory uncertainty is considered, the purely empirical model implemented in OpenQuake includes epistemic uncertainty. For the latter, a weighted source model logic tree is defined considering a variation in the previously derived smoothed seismicity model. For the regional b-value, three branches are defined to accommodate the uncertainty observed from the b-value estimation, the branches being the mean ($b = 1.12$, 40% probability), and +/- one standard deviation ($b = 1.08$, and $b = 1.16$, 30% probability each). In an analogous manner, the maximum magnitude is varied with $M_w 8 \pm 0.25$. Finally, the kernel width for the spatial smoothing is varied with equal probability, using 20 km (average uncertainty as obtained from duplicates during the catalog creation, see Table A.2),

15 km, and 25 km (the latter two considering a further uncertainty of 25%).

In comparison to recorded ground motions, macroseismic intensity observations are numerous for the region (675 observations for 82 events were collected for this study, Haas et al. 2016c). Thus, for the purely empirical model, a ground motion logic tree is constructed applying an optimized selection strategy as proposed in Haas et al. (2016b, , see Appendix A.2) to 10 intensity prediction equations (IPEs; Table 7.2). Please note that in general, the assignment of macroseismic intensity, especially for older sources, leaves room for interpretation, and uncertainties thus become large. Furthermore, note that for calculation and visualization purposes, macroseismic intensity values are treated as decimal numbers throughout this section. These are not to be miss-interpreted as actual values of macroseismic intensity, which are defined as ordinals (e.g., 5.9 and 6.3 are both intensity VI).

Table 7.2.: Considered Intensity Prediction Equations (IPEs), their performance for three measures, similarity cluster and weights¹

ID	IPE	RMSR	LLH	EDR	Cluster ²	SS ³	W ⁴
2	Bakun (2006)	1.32 (5) ⁵	-4.28 (6)	1.96 (7)	a	791	0
5	Dowrick and Rhoades (2005)	1.11 (2)	-4.90 (9)	1.44 (3)	a	791	0
6	Sørensen et al. (2009)	1.00 (1)	-2.20 (1)	1.45 (4)	a	791	0.25
9	Sørensen et al. (2010)	1.51 (8)	-3.03 (3)	2.31 (8)	a	791	0
10	Gasperini et al. (2010)	1.22 (3)	-2.87 (2)	1.37 (2)	b	230	0.25
8	Pasolini et al. (2008)	1.23 (4)	-3.08 (4)	1.50 (5)	b	230	0
4	Chandler and Lam (2002)	1.38 (6)	-5.85 (10)	1.77 (6)	c	0	0.25
1	Allen et al. (2012)	1.76 (10)	-4.13 (5)	3.33 (10)	d	53	0
3	Bindi et al. (2011)	1.60 (9)	-4.49 (8)	2.93 (9)	d	53	0.25
7	Le Goff et al. (2014)	1.45 (7)	-4.46 (7)	1.36 (1)	e	0	0

¹ RMSR, root-mean-square residuals; LLH, log-likelihood method (Scherbaum et al. 2009); EDR, Euclidean distance ranking (Kale and Akkar 2013).

² Similarity cluster the IPE belongs to (see Figure A.1).

³ Within-cluster sum of squares (SS).

⁴ Assigned weight in the logic-tree framework of the hazard model.

⁵ Rank of the IPE is given in braces for each measure.

First, a set of 10 IPEs (Table 7.2) is evaluated over a discretized parameter space with magnitudes varying from M_w 5 to M_w 7 (0.1 steps) and epicentral distances ranging from 5 to 200 km in 5 km steps. Applying the k-means clustering approach to the resulting 1'040 mean estimates with corresponding standard deviations and iteratively minimizing the cluster sum of the square residuals, yields 5 clusters, as shown in Table 7.2. Please note that macroseismic intensity, unlike PGA, is normally distributed and fully described by its mean and standard deviation.

Unfortunately, for most events, source mechanisms are unavailable, therefore only epicentral and hypocentral distances are used. Missing depths were set to 10 km, the median depth observed in the catalog. The performance is evaluated using three simple measures, the root-mean-square residuals (RMSR), the log-likelihood method (LLH; Scherbaum et al. 2009) and the Euclidean distance ranking (EDR; Kale and Akkar 2013). While the LLH is computed from the log-likelihood of the

observations given the respective IPE, the EDR considers the Euclidean distance between predicted and observed intensity, and accounts for a potential systematic first order model bias. For all of the three measures, a value closer to zero stands for better performance. Each of the measures separately would result in very different rankings, e.g., Dowrick and Rhoades (2005) RMSR:2, LLH:9, EDR:3 or Sørensen et al. (2010) RMSR:8, LLH:3, EDR:8 (see Table 7.2). Following the approach proposed the similarity between IPEs is identified by the clustering and thus only one IPE per cluster has to be chosen. Based on the ranking, IPEs are chosen for the logic tree of the macroseismic intensity based models. For the first cluster (a) it is Sørensen et al. (2009, only EDR ranks it second in the cluster). Cluster (b) is best represented by Gasperini et al. (2010) and cluster (d) by Bindi et al. (2011) as only LLH prefers Allen et al. (2012). Regarding the clusters with only a single IPE (c and d), only Chandler and Lam (2002) is considered as, although Le Goff et al. (2014) fits smaller magnitude events well (not shown here, but can be seen from EDR:1), it is applicable only to $M_w \leq 6.2$. As shown in Figure A.1, it strongly overestimates the intensity for large magnitudes. Within the logic tree, each of the selected GMPEs is equally weighted as potential weights that can be obtained, e.g., from the likelihood are not necessarily reliable if data is not sufficient.

As macroseismic intensity is used here no site-amplification is considered because these phenomena are already included in the intensity observation and cannot be clearly separated.

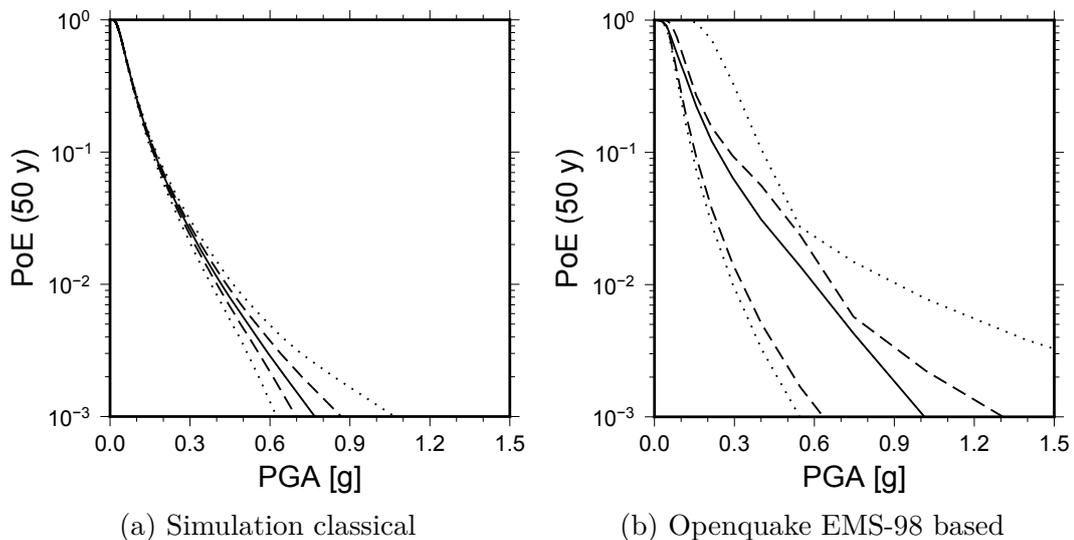


Figure 7.5.: Hazard curves obtained from the EMS-98 based empirical model converted to PGA using Worden et al. (2012) for a single site in central Nablus (35.260 N, 32.225 E) compared to the mean hazard curve obtained from a Poissonian seismicity model using the simulation framework and averaging over all sites considered within Nablus. Quantiles are indicated as: dotted (5/95), dashed (25/75) and solid (50%)

Figure 7.5 shows the resulting mean hazard curve estimated for Nablus compared to the hazard curve for the Poissonian model obtained with the simulation framework converted to macroseismic intensity using the relation defined by Worden et al.

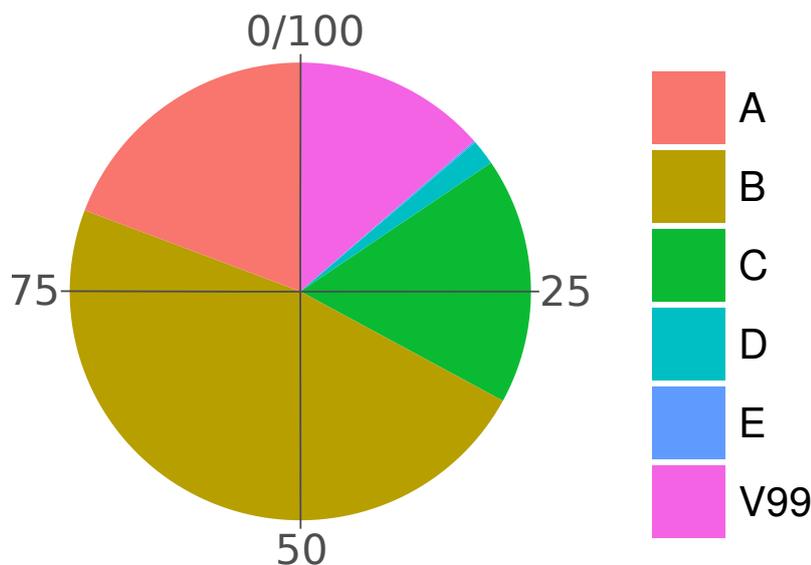


Figure 7.6.: Percentage of EMS-98 vulnerability classes assigned by surveyors to the buildings screened in Nablus using the RRVS (V99: unknown vulnerability)

(2012) and averaging over all sites. The first observation is that including epistemic uncertainty the hazard estimates show a large variability. Furthermore, the median estimates obtained from the simulation framework are lower than the ones obtained from the empirical model. The difference is between 25 and 30 %, i.e., 0.17 vs 0.24 g at 10% PoE in 50%, 0.42 vs 0.6 g at 1%, and 0.76 vs 1.02 g at 0.1%. This is not unexpected, and the reasons for this can be various, for example the GMPE selected in the simulation framework might be underestimating the hazard, but also the conversion from macroseismic intensity to PGA might inflate the empirical hazard.

During the collection of the building data in Nablus the engineers were also asked to assign, based on their experience, EMS-98 vulnerability classes to the screened buildings. Figure 7.6 shows the resulting vulnerability class distribution for Nablus.

Using the same spatial distribution of buildings as for the exposure model used in the simulation framework, but assigning vulnerability classes and fragilities based on Giovinazzi and Lagomarsino (2004) allows one to run a classical risk analysis in OpenQuake for a 50 years period (Pagani et al. 2014).

Table 7.3 compares the average damage transition estimates obtained with the simulation framework using classical risk assumptions ($FR(1)$ in Table 6.1) and the empirical model obtained using the EMS-98 model built for OpenQuake. Please note, that two differed sets of damage states are used: EMS-98 and HAZUS. Thus as with the fragility model comparison, only the first three can be directly compared (see section 7.4). On average, the estimates from the simulation framework seem to underestimate the damages (see Table 7.3). This is expected due mainly to two reasons: 1) the simulation framework yielded a lower hazard and 2) the fragility models used in the simulations are relatively strong. For the directly comparable damage grades 1-3, one observes that while the first damage grade, i.e., damage in the linear

Table 7.3.: Average number of observed damage states within 50 years for the empirical model built in OpenQuake and a model under classical assumptions from the framework ($FR(1)$ in Table 6.1)

EMS-98 Damage grade	1	2	3	4	5
Empirical model	280.82	161.34	66.63	14.87	0.82
HAZUS Damage grade	1	2	3	-	4
Simulation $FR(1)$	313.57	47.60	12.34	-	6.48

part of the damage process, matches well (overestimate of about 12%), the higher damage states 2 and 3 don't agree well. For the second damage state, the simulation framework yields only about 30% of the empirical estimate and for the third damage state, about 19%. Although, the highest damage states are not directly comparable, they seem to be around 50% for the simulation results with respect to the empirical model. One can conclude that most likely as assumed already, previously presented absolute loss estimates underestimated the actual risk. Nevertheless, the results can be used to draw some conclusions relevant for future research. The relevance of the presented results and under what limitations reasonable conclusions can be drawn, is briefly discussed in the next chapter.

8. Relevance of the findings and limitations

In the light of the previous validation exercises, it is clear that the results obtained within this thesis can show only a small facet of the problem, and cannot be necessarily generalized to their full extent. Although absolute losses are not robust (see Table 7.3), the observed tendencies among the different models are defensible and could all be attributed to causes as discussed in Chapter 6. The observed tendencies (Figures 6.1 and 6.5) fall within the range given in previous studies, for example, 25% loss increase for the lifetime loss of a single structure considering fore- and aftershocks as estimated by Yeo and Cornell (2009b), and the 207% maximum increase in the probability for the highest damage level expressed as 60% period elongation observed by Trevelopoulos and Guéguen (2016) for Greek building typologies.

Despite the promising results, several limitations have to be considered before drawing conclusions from the results. For each of the components, i.e., hazard, exposure, damage and the repair process, simplified assumptions were made. The hazard component currently considers only a single Gutenberg Richter b -value, which is an essential component of the seismicity model and might have a strong influence on the results. Considering the latter, while this is true, as was shown in Haas et al. (2016b) for the macroseismic-intensity based Poissonian hazard model, the influence of this factor is dwarfed by the influence of the chosen ground motion attenuation model. For the ground motion model, a consistent selection of these models, such as the process proposed in (Haas et al. 2016b, see also Appendix A.2) and demonstrated for the macroseismic-intensity based model (Haas et al. 2016b), is essential for reliable risk estimates. As only a single GMPE was used within the simulations (see section 5.1.2), it is clear that the presented results are in reality subject to much larger uncertainties than observed from the limited models here. Using PGA as a ground motion parameter for the analysis is not the optimal choice, as other parameters such as SA are better correlated with building response, especially for larger buildings. This choice was made mainly to estimate only a single parameter in the ground motion calculation instead of several per site, as would be required for SA. This can be improved in a future revision of the framework. Regarding the occurrence of aftershocks, considering only an isotropic spatial probability model around the triggering event (see equation 4.7) is another limiting factor, as aftershocks follow the present seismic structures and are not arbitrarily spaced around the triggering event. Considering additional earthquake parameters, such as rupture geometry or stress drop, could yield a more realistic spatial distribution (see e.g., Bach and Hainzl 2012). Another strongly limiting factor regarding the hazard component is the consideration of site-effects only through an approximation from

topography (see Figure 5.5). While the usage of v_{s30} to approximate site effects has limited predictive power, as it is only partly correlated with site amplification, the usage of such inferred values is potentially deceptive and strictly speaking not suitable for urban scale analyses. Nevertheless, more reliable options were not considered, as their determination requires a thorough analysis, which is beyond the scope of this thesis (see e.g., Parolai 2012). In particular, for the latter reasons results obtained within this study are aggregated to the whole town (Figure 6.1) and not differentiated by the spatial distribution of assets, although the framework would allow one to do so. An actual application of the presented framework beyond demonstration purposes should not ignore the spatial differentiation, as this, as described in the introduction, is paramount information for disaster risk reduction in terms of urban planning and emergency response.

The framework currently considers only aleatory uncertainty in the seismicity and ground motion simulation. However, although not shown here, epistemic uncertainty through the usage of a logic tree could be introduced quite easily. Regarding the consideration of the exposure in the model, four main limitations remain. For the presented simulations, only an approximately 10% sample of the full model was used. Furthermore, the building type distribution was not differentiated spatially, as the derived strata would have allowed to do. In addition to the limited consideration of site-effects as mentioned above, this decision is mainly based on the relatively low number of only 827 screened buildings (see Figure 5.10) and the desire for a statistically reliable model. In addition, no uncertainty was considered in the exposure model and only a static (time-invariant) model was used (see Figure 5.14).

The component with the most serious limitations within this framework is the fragility models. First of all, the link between the fragility and exposure models is relatively weak, as only a visual screening was performed during the exposure analysis, with analytical fragility models needing to be derived to consider cumulative damage effects. Fragility models were obtained using a strongly simplified approach and only restricted parametric variations were considered during their derivation (see Appendix C.1 and C.4). Within the current framework, this parameter uncertainty in the fragility model is dealt with by sampling the distribution of fragilities for a building type only once per building type, rather than for each individual building. Although they comprise only about 3% of the building stock (see Figure 5.12), a reliable model for Nablus should also include the unreinforced masonry structures within the town. Furthermore, the abundant building irregularities in Nablus have not been considered. This is mainly due to the fact that most of these irregularities are related to torsional eccentricity, which cannot be modeled with SDOF systems. Only the very common presence of soft stories in the ground floor could be incorporated with relative ease in the applied approach, but it was not for reasons of consistency. Regarding the repair models employed here (Figure 4.1), the main limitation is that these depend only on time and thus hardly capture the complex decision-making process commonly involved when repairing structures. A further limitation regarding this is that the simulation framework currently lacks functionality to consider partial repair, retrofitting or replacement with stronger buildings.

Regarding the economic losses, the frame work is currently not estimating losses due to business interruption or building content, which may exceed the structural and non-structural losses. However, at least the former could be introduced quite easily, as repair times are recorded during the simulation. For the single year simulations (section 6.4), two limiting factors have to be considered for such a short simulation period: 1) one year simulations are strongly conditioned by the input catalog that is used to model sequences resulting from events prior to the observation period, and 2) simulating only a single year may cut seismic sequences that may extend over several years. In addition to these, only a small number of simulations were used that resulted in strong undersampling at low PoEs (Figure 6.15). The longer period simulations (Figure 6.5) are thus considered to be more reliable.

9. Conclusion and Outlook

The first aim of this thesis was to develop and test an approach that allows the construction of time- and state-dependent risk models over urban scales. This was realized using a simple simulation framework that integrates a number of components from previous studies, i.e., an ETAS model as described by Beauval et al. (2006), ground motion calculators provided by Pagani et al. (2014), a Markov-Chain damage model as proposed by Yeo and Cornell (2005), and state-dependent fragilities and time dependent repair models developed herein. All the individual components were integrated into a single consistent framework and implemented as a computer program. Testing the approach on the city of Nablus yielded time- and state-dependent risk estimates. For the simulations, two observation periods were considered: first, a long observation period of 50 years, which is relevant mainly for urban planning and disaster risk preparedness, and second, annual risk estimates, which are more commonly used in the insurance industry. As some of the model components developed within the application are only intended for demonstration purposes and were thus simplified, the resulting absolute losses are not accurate. This was discussed in the validation chapter and it is concluded that despite the limited accuracy, the models are precise enough to be judged in terms of general tendencies.

Employing the modularity of the framework, the second objective of the thesis, i.e., to compare the results to a model employing classical estimation schemes, could be reached. Differences were tested for significance using a two-sided Kolmogorov-Smirnov test and a significance level of $\alpha = 0.05$. Considering time- and state-dependency in seismic risk assessments over urban scales and a observation period of 50 years, yielded a **significant difference in the estimated losses**. When compared to a classical model, the following differences were noted:

- the median hazard increased by 20.5% at 1%PoE (Figure 5.6),
- the median losses increased by 58.4% at 1%PoE (Figure 6.1),
- the difference between both models increases exponentially with lower PoEs (Figure 6.2),

Note that the last point is mainly due to aleatory uncertainty (except for the small epistemic uncertainty in the building models in Table 5.7). These findings agree with previous studies on individual structures (Yeo and Cornell 2005; Yeo and Cornell 2009b; Iervolino et al. 2016; Trevlopoulos and Guéguen 2016; Tesfamariam and Goda 2017), and regional scales (Papadopoulos and Bazzurro 2018), and show that

trends observed therein are replicated over urban scales. One can thus conclude that time- and state-dependency in seismic risk assessments is not negligible over any scale.

The third research question of this thesis aimed to determine the relative influence of individual model components and decisions. This was addressed performing additional simulations for which the model components and configurations were modified accordingly and several characteristics could be identified. In general, the trends were found to vary with the PoE by 2-5% for the individual models (Figure 6.5). Among all the models tested, the ones considering full seismic sequences showed consistently higher losses over the full PoE range (Figure 6.5).

Simulations that introduced delayed-repair and state-dependency within a **Poissonian seismicity model** showed no pronounced increase in losses when compared to a classical model (Figure 6.7). The small observed differences were found to be non-significant according to the Kolmogorov-Smirnov test ($\alpha = 0.05$; Table 6.3). Thus, an artificial double counting effect, as in the case when assuming immediate repair within an ETAS model, was not observed for the classical model. Such an effect would show a tendency towards lower estimates at low PoEs for the Poissonian seismicity model with a delayed-repair model and state dependency, but this was not observed due to sufficiently long inter-event times (Figure 6.7). This, however, can be a result of the model conditions, i.e., underestimated fragility and only moderately high seismic hazard. Thus, these factors might not be irrelevant in cases considering different model conditions.

As was shown in Figure 6.8, the influence of the **shape of a lognormal repair model** is rather complex. When compared to the conservative model, using the *optimistic* repair model resulted in slightly higher losses (around 2%) with respect to a model with the *conservative* repair assumption. As shown in Table 6.2, this was found to result from:

- fewer transitions from damaged states (cumulative effect),
- but more transitions from the intact state

and thus again due to the previously observed double counting. However the Kolmogorov-Smirnov test showed no significance for this difference ($\alpha = 0.05$; Table 6.3). In contrast, the slightly lower (by up to 5%) losses when comparing estimates employing a *pessimistic* model to estimates of one with the *conservative* model were found to be significant. In this case the difference stems from (Table 6.2):

- more transitions from damaged states (cumulative effect),
- lower number from the intact state,

and thus the buildings stay longer in high damage states and are not damaged again. In general, the variations resulting from the changes in the shape of the

lognormal repair model were small. However, these observations are limited in the sense that only rarely are buildings not retrofitted when they have been damaged, or not replaced by more resilient structures when destroyed.

Comparing the means of the loss distributions at various PoE levels found the largest average size of effect (Cohen's d with pooled standard deviations) for the full model with conservative repair and the classical risk estimate ($d = 3.09$ at 1% PoE and $d = 5.88$ at 10% PoE, Figure 6.11). The average lowest, but still significant, size of effect on the mean of the loss distributions was observed for the models with *pessimistic* and *optimistic* repair options ($d = 0.13$ at 10% PoE and 0.21 at 2% PoE; Table F.1). The difference in the mean between a classical model and a classical with delayed repair at 5% PoE (Figure 6.11) was found in general to be the lowest significant size of effect ($d = 0.03$).

The influence of the **observation period** was evaluated by distinguishing only between single year and 50 year simulations. Due to the lower support when compared to simulations using the longer simulation period, the estimates from annual simulations (25'000 years of seismicity instead of 500'000) showed stronger variations (Figure 6.15). However, both showed the same tendencies (Figure 6.13) and a rapidly increasing difference (Figure 6.14) if the time- and state-dependent estimates are compared to classical estimates.

The last research question of this thesis considered if simplifications can be made. For the 50 year simulation, the question of weather sufficiently accurate risk estimates can be obtained by modeling full seismic sequences, but **dropping state-dependency or generally neglecting delayed-repair**, was investigated.

Using simulations with long observation periods, the influence of the delayed repair, state-dependency and the shape of the time-dependent repair model could be tested. For the **delayed repair and state dependency**, neglecting either yields similar losses above 10% PoE, but observed differences were found to affect losses significantly. With respect to the full model, neglecting state-dependency, but modeling delayed repair and full sequences increases loss estimates with respect to a classical model (Figure 6.5). However, this increase is between 0 and 28% less pronounced than that observed for the full model. The difference between the full model and such a **state-independent model** is inversely correlated with the PoEs, i.e., increases for lower PoEs (Figure 6.5).

In the case of **neglecting both** delayed-repair, and thus also state-dependency, the estimates agree with those of a full model at high PoEs, but start to strongly overestimate the losses below 10% PoE, due to the artificially higher number of low-level damage transitions obtained for sequences with several strong events (Figure 6.5, Table 6.2, and Figure 6.6).

For annual seismic risk models, the question is if a time-variant repair model is necessary, or can be either replaced by assuming no repair to happen within a year or assuming a constant (fixed) probability of daily repair, was investigated. For both cases, no strong deviation from the more complex time variant model was found for

PoE $> 0.2\%$ (Figure 6.15). However, the small difference between estimates from a model with time-variant repair and one with no repair was tested to be significant ($\alpha = 0.05$) over testable PoE ranges, albeit with the lowest significant size of effect ($d = 0.02$ until 0.22). According to the same test, a repair model with a fixed constant probability model ($0.000623/\text{day}$) was found to replicate the estimates of the more complex model sufficiently (Figure 6.15), i.e., yielded no significant differences in the testable PoE range (Table 6.5). However, tendencies observed for low PoEs might indicate a difference at lower levels than could be tested with the number of simulations used here (Figure 6.16).

The question of whether a Poissonian model used to trigger sequences in the ETAS model, is sufficient to obtain comparable classical risk estimates, or if actual main shocks have to be extracted, was also investigated. Using an annual seismicity model, it could be shown that using the simpler approach yields estimates that are mostly only slightly below a real mainshock model (except for a small peak observed in all ETAS based models, Figure 6.15), and no significant difference for the testable range of PoEs could be detected (Table 6.5).

While most of the conclusions above have general implications, however, note that the accuracy of the models used is limited and the size of the differences might vary depending on the components. For accurate estimates, more thoroughly developed hazard, exposure, and fragility models than those employed for demonstration purposes here are required. In general, one can expect the effects to be more pronounced than has been shown in this work.

Concluding from the results of this study, the large efforts involved in determining time- and state-dependent seismic risk models rather than classical Poissonian models can be justified, considering the strong impact of both effects on the resulting seismic risk estimates. The recommendations that can be drawn from the presented results are that despite the greater efforts required, time- and state-dependency should be considered for any seismic risk assessment. Considering full seismic sequences, but no delayed repair should be avoided, but in case of short simulation periods and PoEs $> 1\%$, the time-variant repair model can be replaced by a constant probability repair model.

The presented framework builds a consistent basis for further research in a multitude of directions. The next research question could be to apply the framework to a situation for which observed data is available, such as, e.g., the recent sequences in Italy in Rieti 2016, or Emilia Romagna 2012. Furthermore, several improvements can be envisaged for individual components. Examples could include considering other mechanisms in the seismicity model that describe seismic sequences, such as the region-time-length (RTL) model (Sobolev and Tyupkin 1997) which is based on observed seismicity within space- and time-windows and considers rupture sizes, inter-event time, and source-to-site distances (Chen and Wu 2006) or using a Markov Chain model (Aki 1956; Vere-Jones 1966). A more reliable site-effect estimation would allow one to consider effects due to the exact location of a structure more consistently and to distinguish these effects between individual structural types. Replacing the strongly simplified fragility models by more realistic models would

yield better insight regarding differences in behavior for various building types, also with respect to events stemming from different tectonic regions (crustal, inter-slab or intra-slab events) such as were found by Tesfamariam and Goda 2017 for British Columbia, Canada.

The current static exposure model could be replaced by a stochastic model that includes dynamic changes, for example, due to changes in occupancy, retrofitting and deterioration, or replacements of structures. Such a model would allow one to consider different policies, e.g., build-back-better, or changes in building codes. Being a simulation framework, even urban growth could be considered, using agent-based models that drive changes in the exposure (Waddell 2002). Such models are often used in urban planning and used to replicate the socio-economic behavior of people, and effects of their decisions on the built environment. As damages are recorded for each building, these could in turn induce a feedback in socio-economic behavior. Using oriented graph networks, effects and interactions for infrastructure and life-lines could be included. For such a case, in addition to losses from damage state transitions, disruption costs based on the recorded times of repair could be obtained. To calculate these for buildings, one would differentiate in the exposure between different occupancy types, at the very least between residential, commercial, and industrial buildings.

Furthermore, the current, only time-dependent, repair models could be extended to time- and state-dependent models, making use of more detailed data, such as is available for the Bam 2003 earthquake (Hashemi-Parast et al. 2017). The strong correlations between damage grade and repairs observed therein could potentially be captured in a more generalized way if enough data is collected. For complete reconstruction, analytical construction estimates that depend on the size or volume of the structure could also be considered (Schneider et al. 2012, 27ff).

Finally, the framework could be extended to consider multiple hazards. If hazards are considered as being independent, only a spatially distributed history and the resulting intensities or hazard levels for the same observation period would need to be simulated. Furthermore, if damage mechanisms are considered to be independent for two hazard types, only a set of fragilities or vulnerability functions for the respective hazard have to be added. In contrast, if the interaction between hazards is to be taken into account, or even the interaction of damages arising from different hazards, this can become very complex and opens a large field of potential research questions that may be answered by extending the presented framework.

The work presented here can thus be considered an early step towards time- and state-dependent risk assessment over urban scales. With the increasing amount of sensors, collected data and also improved computational capabilities of modern computing architectures, using, e.g., distributed and cloud computing, the manageable complexity of simulations will increase dramatically in the near future. This will allow more physics-based model components to be developed, that can replace the still abundant empirical models within risk assessments. Extending simulation frameworks such as the one presented here and combining it with a complex city exposure model opens the way to simulate an actual *City at Risk*. Figure 9.1 shows



Figure 9.1.: City at risk: Combining sophisticated socio-economic models and time-dependent hazard simulations will yield a model suitable to explore interactions on the hazard and risk levels. Agent-based models can simulate the dynamics within a city (Waddell 2002), where people are moving, buildings are replaced or new once erected, whereas advanced hazard simulators can be used to model their impact on the city. Finally, a frame work that controls the interactions among all components will allow one to obtain a time- and state-dependent multi risk model.

an idealized sketch of such a frame work which could consider the dynamics of the population, demographic changes, and how these shape the very place the people live in. Having such a complex dynamic system and modeling potentially disastrous natural phenomena such as earthquakes, floods, volcanoes, or storms would allow one to explore the interactions not only among hazards, but also from a risk perspective. If considering probabilities rigorously, or at least consistently, this could highlight potential vulnerabilities and help exploring the risk continuum beyond the limits of our experience.

Appendices

A. Hazard related

A.1. Seismic catalog

A seismic catalog for the Dead Sea region was derived from several sources (see Table A.1; Haas et al. 2016c; Haas et al. 2016b).

Table A.1.: Composition and Sources of the Declustered Catalog.

ID ¹	Source ²	Events	Period	Mw Range	Missing z ³
1	Agnon (2014)	3	363–659	6.0–6.8	3
2	Marco et al. (2003)	1	749–749	7.3–7.3	1
3	Avni et al. (2002)	1	1927–1927	6.3–6.3	0
4	Sbeinati et al. (2005)	24	37–1822	6.1–7.5	1
5	Ambraseys (2006)	10	551–1872	6.0–7.2	5
6	Ambraseys and Adams (1993)	33	1894–1986	4.5–6.0	12
7	Grünthal and Wahlström (2012)	732	1002–2006	4.0–7.5	145
8	GII	696	1905–2014	4.0–6.1	0
9	NCEDC b	476	1963–2014	4.0–6.1	0
10	ISC-GEM	2	1905–1914	5.6–7.0	0
11	IRIS	1268	1965–2014	4.0–6.1	2
12	ISC	2	1979–1984	5.6–7.0	0
13	KOERI	95	1903–2014	4.0–7.1	1
14	NCEDC a	9	1955–1994	4.0–5.7	0
15	Giardini et al. (2013)	20	1003–1997	4.1–6.8	19
16	GSHAP	10	23–1940	5.5–6.7	7
Total	3382	23–2014	4.0–7.5	196	

¹ The ID represents the hierarchy (i.e., in case of duplicate events), that is which catalog was preferred over another, respecting the defined thresholds (see the Historical Seismicity section).

² GII, Earthquake in Israel bulletin; GSHAP, Seismic Catalog of Turkey and Neighbouring Countries of the Global Seismic Hazard Assessment Programme; IRIS.: FDSNWS-Event database; ISC, The Bulletin of the International Seismological Centre; ISC-GEM, ISC-GEM Global Instrumental Earthquake Catalogue v.1.05; KOERI, National Earthquake Monitoring Center EQ Catalog; NCEDC, ANSS Comprehensive Catalogue accessed from two different query builders (a) and (b) (for details on all these data sources, see Haas et al. 2016c; Haas et al. 2016b)

³ Number of events without depth estimates.

Figure 5.1 shows the resulting catalog. Unfortunately, most catalogs do not provide uncertainties with the respective estimates. Thus, a quantitative comparison of

Table A.2.: Time and distance thresholds, that if not exceeded identify duplicate events in the catalog.

Period	Time	Distance ($^{\circ}$)	Argument
Before 1000 C.E.	1 day	0.5	Many records of events provide only dates
1000–1899 C.E.	12 hours	0.4	Most records of events provide approximate times of the day
1900–1959	1 hour	0.3	Well documented
Since 1960	2 min	0.2	Instrumental catalogs

the different catalogs' quality is hindered and a subjective hierarchy given in Table A.1 has been followed to define a preference in case of duplicates. Duplicates are identified using subjective time and distance thresholds (A.2) and giving priority to complete hypocenter entries. Assuming the improvement of parameter coherence among catalogs over time, the time dependent thresholds for time and location precision, the catalog is divided into the period before 1000 C.E., the period until 1900, the early instrumental period since 1900, and the instrumental period since 1960.

The hierarchy among the catalogs (ID in Table A.1) is chosen based on the following assumptions: the most reliable estimates should stem from sources focusing on single or only few historic events (ID 1-6), followed by one of the longest (1000 till 2006 C.E.), most reliable and complete available catalogs for the region, the European Mediterranean Earthquake Catalogue (EMEC; ID 7; Grünthal and Wahlström 2012), followed by instrumental catalogs where the highest rank is given to the local catalog of the Geophysical Institute of Israel (ID 8), followed by other global sources (ID 9-11), complemented with both purely instrumental catalogs (ID 12,13), and the Seismic Hazard Harmonization in Europe (SHARE) project catalog (SHEEC; Grünthal et al. 2013; Stucchi et al. 2013), as well as the Global Seismic Hazard Assessment Program (GSHAP) Turkey catalog (ID 14,15). Details on the data sources are available in Haas et al. (2016c) and Haas et al. (2016b). Events that occurred B.C.E. are excluded as these are very uncertain. If available, the moment magnitude (M_w) of the events is used, if not, with decreasing preference, surface magnitude (M_s), local magnitude (M_L), body wave magnitude (M_b), or duration magnitude (M_d , M_c) estimates are employed. Events with unknown magnitude type or zero depth were excluded. The magnitudes are then harmonized to moment magnitude M_w , using different conversion schemes, that have been identified as being well suited for the eastern Mediterranean by Çağnan and Kalafat (2012) (see Table A.3).

The resulting final catalog, includes 14'510 events above M_w 3.

Table A.3.: Magnitude conversion relations applied to the original magnitude types in the catalog.

Magnitude	Conversion	SDR (M_w) ¹
M_S	Scordilis (2006)	$0.17(M_S \leq 6.1); 0.2(M_S \geq 6.2)$
M_L	Akkar et al. (2010)	0.2
M_b	Johnston (1996)	0.26
M_D/M_c	Çağnan and Kalafat (2012)	0.3

¹ Standard deviation of the residuals (SDR) determined by the authors of the respective relations.

A.2. Optimized selection of ground motion models

Following Haas et al. (2016b), a simple two-stage approach motivated by Scherbaum et al. (2010) is proposed, which can support the selection of ground motion models within a logic-tree framework, that is less biased with respect to the criteria as defined by Bommer and Scherbaum (2008). The fundamental idea is to select attenuation models that are not too similar, but still perform well with respect to the observed data. In order to make the available models comparable, one can evaluate each over a discrete parameter space, at least magnitude and distance, which yields comparable sets of ground motion estimates. This dataset can then be grouped by similarity. For this purpose, a clustering method is proposed (Haas et al. 2016b) to identify similar GMPEs and reduce their number to be included in a realistic model. Clustering is the mathematical procedure to find groups in data, where each element within its group is mathematically more similar to other members of the group than to elements of other groups (Kaufman and Rousseeuw 2005). Several different algorithms and also different measures of similarity exist for this purpose. A simple and popular algorithm is k-means clustering (Steinhaus 1956) which separates a dataset into a specified number of k groups. For this purpose, k points are defined in the similarity space which can be, e.g., euclidean distance, and each element is assigned to the respective closest point, forming clusters. Minimizing the distances one obtains optimal points defining centroids for each cluster (Kaufman and Rousseeuw 2005, p. 41). A critical choice is the optimal number of clusters k and usually found in an iterative way, e.g., minimizing the within cluster and total sum of variances. Applying this to a data set, obtained by evaluating all GMPEs over the discrete space, yields clusters of similar models.

Once all models are evaluated over the discrete parameter space, a tool such as Sammon's map 1969 can be used to visualize the similarity of multi-variate models such as IPEs or GMPEs in two dimensions, as also described in Scherbaum et al. (2010). Using a distance metric such as the euclidean or Kulback-Leibler distance (as proposed by Scherbaum et al. 2010), the distance between the model estimates (e.g., vectors of mean intensity) obtained from the discretization, can be mapped into a two dimensional space and a Sammon's map as in Figure A.1 can be generated. Here 10 IPEs (Table 7.2) have been selected as potential candidates to model macroseismic intensity within the Dead Sea region (Haas et al. 2016b), the distance between the models corresponds to the Euclidean distance between the predictor vectors for the means. Figure A.1 thus can be used to visually identify the grouping resulting from the k-means clustering.

Figure A.1 b) – f) exemplifies for each of the clusters how well the IPE estimates for a M_w 7.25 event at 10 km depth fit data for regional events with magnitudes between M_w 7 and 7.5.

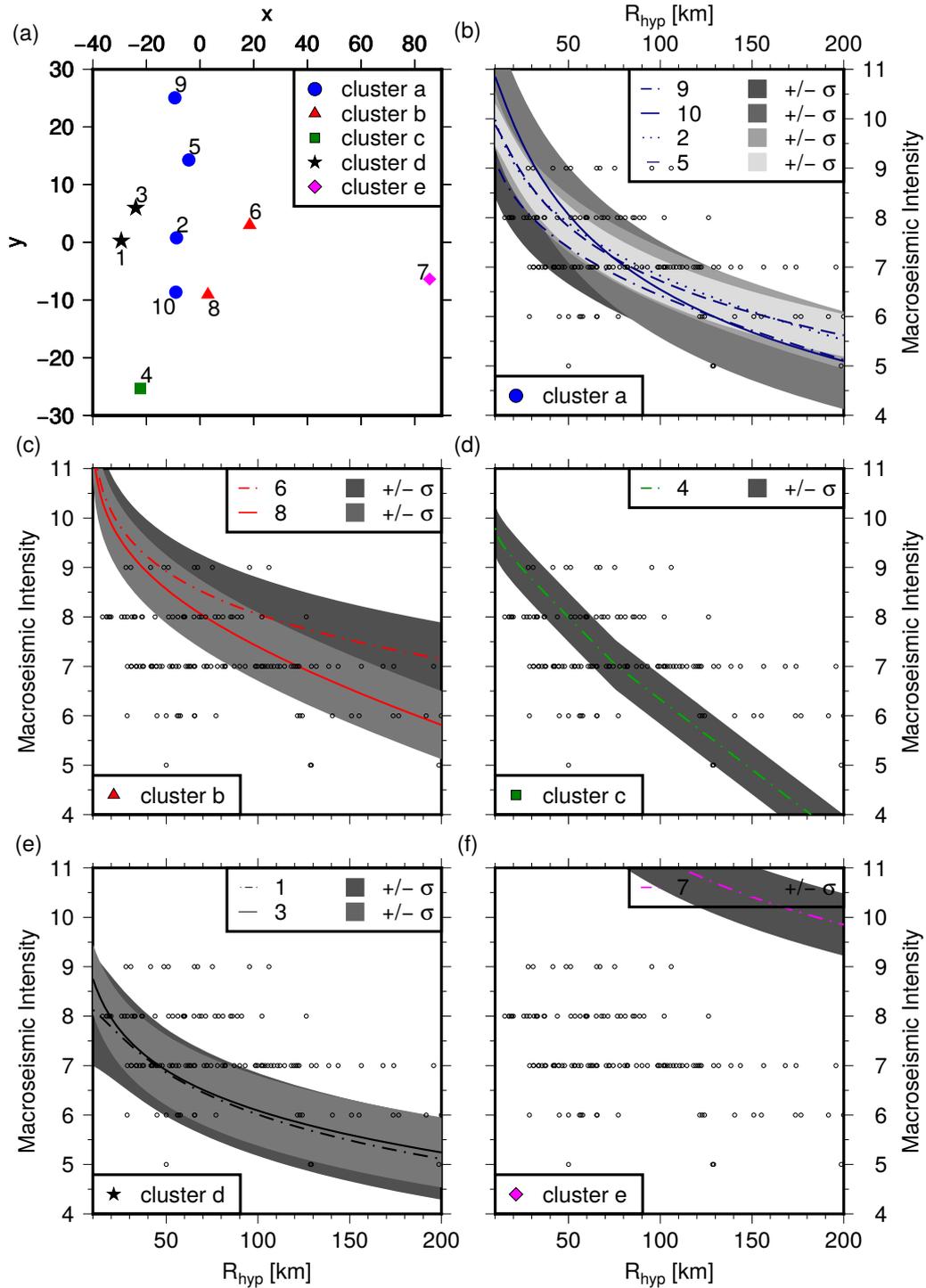


Figure A.1.: (a) Sammon's map (Sammon 1969) of the 10 considered intensity-prediction equations (IPEs); the symbols indicate the cluster that each IPE belongs to, according to a k-means clustering (Steinhaus 1956) of the mean and standard deviations of their estimates over a discrete parameter space. (b)-(f) Estimated mean (curves) and 1σ (shaded) intensity attenuation resulting from each IPE in each cluster for an M_w 7.25 event at 10 km depth against historical distance intensity pairs for events with M_w 7.0–7.5. IPE IDs are (see also Table 7.2) 1, Allen et al. (2012); 2, Bakun (2006); 3, Bindi et al. (2011); 4, Chandler and Lam (2002); 5, Dowrick and Rhoades (2005); 6, Gasperini et al. (2010); 7, Le Goff et al. (2014); 8, Pasolini et al. (2008); 9, Sørensen et al. (2009); 10, Sørensen et al. (2010). Figure modified from Haas et al. (2016b)

In a second step, one or more standard ranking methods can be used to rank the models within each of these clusters. The best ranking model in each of these clusters can then be chosen for implementation as a single representative for this type of ground motion models. This approach helps to include models in logic tree branches which are mutually exclusive and collectively exhaustive. What remains is the problem of assigning a probability to each of these branches. With the introduction of their ranking schema based on log likelihood, Scherbaum et al. (2009) suggest assigning probabilistic weights to the branches using the respective likelihood. Unfortunately, this is only mathematically valid if hypothetically an infinite number of observations is available (Scherbaum et al. 2009). Thus, assigning weights by expert judgment is often inevitable.

B. Exposure related

B.1. Simplified satellite analysis

Although individual buildings are not captured in medium resolution satellite imagery the level of detail of information that can be extracted is remarkably high. This is due to the multi-spectral content of the imagery and the possibility to derive high level image features such as textural content, if images are segmented (Wieland et al. 2012). In order to divide an urban satellite scene into homogeneous strata several approaches can be followed ranging from unsupervised clustering approaches working on the single image pixel values to supervised learning methods such as support vector machines (SVM; Boser et al. 1992) which can work on image objects such as segments. The main task of a SVM is to predict class labels for unseen data according to their attributes and based on the knowledge obtained from attributes of a training data set (Hsu et al. 2003). Unfortunately, gaining high-level information requires well defined ground truth data, i.e., detailed and verified information to train the SVM that performs the image segmentation and classification task. In many cases, such data is not available and the only visually derivable information may come from high-resolution data such as available through Google Maps (Google, DigitalGlobe 2016) or Bing (Microsoft, DigitalGlobe 2016). Depending on skill, but also knowledge about the region of interest, an operator deriving such data might only be able to distinguish basic categories of homogeneous regions within a town. These could, for example, be distinguishable in terms of dense and sparsely populated residential areas, commercial areas, different types of land use such as agriculture, bare soil, forest or water bodies. For such a level of detail, the sophisticated segmentation and classification process employed by Wieland et al. (2012) might be unnecessary. Simple pixel-based supervised classification approaches might be sufficient instead of the superior, but more complex approach. In comparison with object based imaging, which allows the derivation of high-level features for individual segments of the image, this limits the analysis to the spectral properties of individual pixels. Following a simple pixel based approach, reliable classifications distinguishing the level of detail as retrievable visually can be obtained. For this purpose, a small tool, called Rapid Environmental Mapping - Satellite Exposure (REM SatEx), for the free and open geographic information system QGIS (QGIS Development Team 2017) has been developed to ease the process of classification using a pixel based SVM. The tool is based on the libsvm library (Chang and Lin 2011) as implemented in the Orfeo Toolbox library (Christophe et al. 2008).

Given a training data set of land use and land cover classes, e.g., in a simple case, say built-up and soil, the algorithm defines the hyper plane that best separates these

two classes in the spectral parameter space of the image pixels within these defined regions (e.g., a polygon). In this simple case, one would distinguish pixels belonging to built-up areas and those belonging to soil. The hyper plane resulting from the optimized separation of each pair is then described through the support vectors defining it. Using these vectors, each pixel of the image is then classified according to its level in each of the spectral bands, i.e., its location within the hyper space. The resulting classified image for such a simple case within only two classes would then be an image where each pixel is assigned either the built-up or soil label.

The performance of such classification algorithms is usually evaluated using a test data set's predicted class labels and compare these to the defined ground truth. In the simplest case contingency matrices comparing produced and reference levels can be used to estimate the accuracy of the resulting classification. An example is shown in Table B.1, where four built-up and six soil pixels were defined in a ground truth data set. Regarding built-up pixels, the SVM assigned three pixels correctly and misinterpreted one, whereas four soil pixels were assigned correctly and two wrongly. Ideally, the off diagonal entries should be close to zero. The training set may be improved by adding new training features and/or removing problematic features, e.g., the bare soil surrounding a farm classified as built-up.

Table B.1.: Simple contingency matrix

		Ref. ¹	
		built-up	soil
Prod. ²	built-up	3	2
	soil	1	4

¹ Ground truth pixels defined in the testing set

² Label produced by the SVM algorithm

If the defined classes are exhaustive, i.e., capture all land use and land cover classes present in the satellite scene, the classification will improve. Once a SVM is trained it can be applied to any region, given the defined classes are general and exhaustive enough. Using a sufficient number of land use and land cover class types where built-up regions are not only identified, but also distinguished among each other based on their features, the approach can yield a stratification of a town into homogeneous regions with relatively small effort. Ideally, a recent satellite scene is used to obtain an up-to-date stratification. In addition, some of the medium-resolution satellite imagery are available reaching back several decades in time. For example, scenes from the Landsat satellite series (Data available from the U.S. Geological Survey) are available as early as the 1970s. While the sensor quality has improved significantly over time, older imagery is in most cases sufficient for distinguishing between built-up and non-built up areas. This allows time-series analysis to be performed and the estimation of a lower boundary of the new development areas and thus construction age of parts of a town built after this period. Combining both information allows a simple stratification of a town to be obtained.

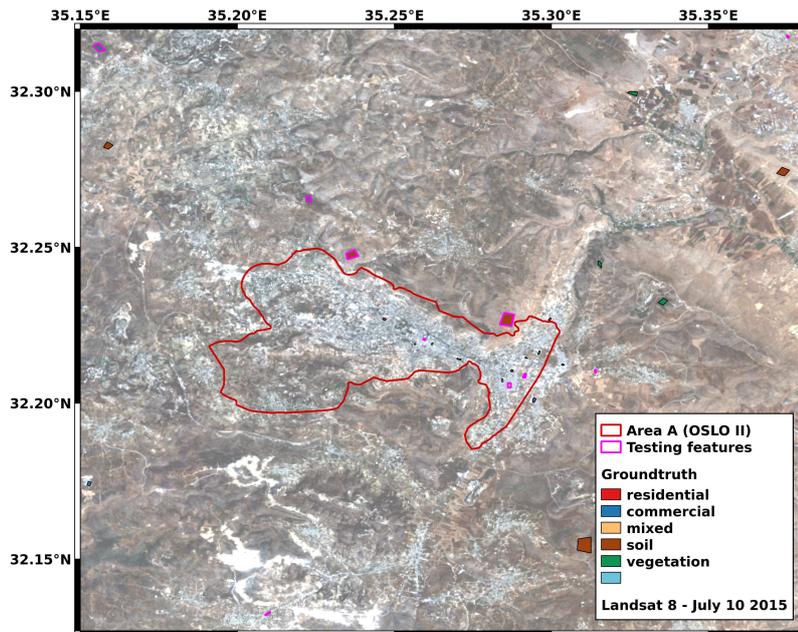


Figure B.1.: Training and testing set derived for Nablus for the five classes residential, commercial, mixed built-up, soil, and vegetation.

B.2. Homogeneous strata for Nablus

A small training and testing data set was derived using high-resolution DigitalGlobe satellite imagery from Google Maps (Google, DigitalGlobe 2016) through the OpenLayers plugin in QGIS, where polygon features delineating different land-use and land-cover classes, namely, residential (5 features), commercial (7) and mixed (6) built-up areas, soil (6), and vegetation (5) are derived. Their size varies from only 2 to 25 Landsat pixels (each 30 by 30 meters). For the purpose of this study, this data set is considered as the ground truth. Of course, the level of detail is limited if no ancillary data is available and the reliability strongly depends on the image quality, as well as the experience of the analyst.

The data set derived for Nablus is shown in Figure B.1. Figure B.2 gives examples for the derived classes, namely residential, commercial, mixed built-up, soil, and vegetation.

Using the simplified satellite analysis on a Landsat 8 scene from the 10th of July 2015 yields a land-use and land-cover classification. Relevant for the purpose here is to successfully distinguish the built-up and non-built-up areas and to categorize built-up areas into more detailed entities.



(a)



(b)



(c)



(d)



(e)

Figure B.2.: Example of training features (red polygons) for (a) residential, (b) commercial, (c) mixed, (d) vegetation, and (e) soil; visually derived from Google satellite imagery (Google, DigitalGlobe, Geoeye 2017).

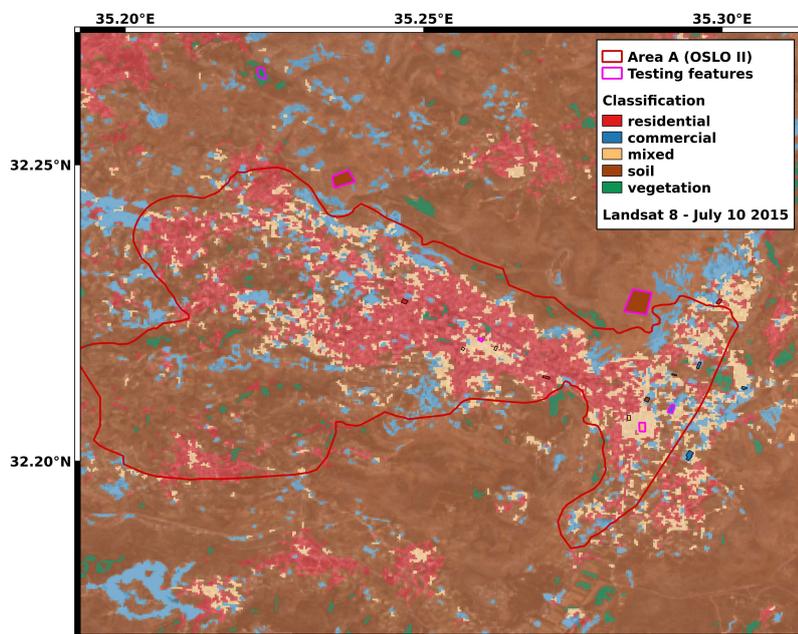


Figure B.3.: Resulting classification using REM SatEx for residential, commercial, mixed built-up, soil, and vegetation classes (see Figures B.2).

B.3. Ground based data sampling

The strata derived from the satellite analysis can be used to drive the prioritization of a ground-based sampling of individual structures. The most efficient strategy is to use a mobile mapping system as proposed by Wieland et al. (2012). A representative sample can be obtained by sampling individual points in the strata and ensuring that each strata is sampled consistently. Ideally, sample sets depend on the extent of each strata. Once the sample points are chosen, an optimal route to reach each of these points with the mobile mapping system needs to be determined. To determine such a route, referred to as a traveling sales man problem, requires detailed information on the street network. A free-of-cost and remarkably complete source for street data is the Open Street Map (OSM; OpenStreetMap contributors 2017). This data is crowd sourced, usually manually digitized from high resolution satellite imagery. As a result, the network may have broken connections or isolated segments. Such deficiencies can lead to problems with routing algorithms which are based on representing the street network as directed graphs. Thorough analysis and in some cases the cleaning of the network topology is thus necessary. A helpful tool for such a task is the topology-based GRASS GIS which provides algorithms to identify such imperfections. Once its topology is flawless, the network can be turned into a routable graph network using PostGIS (Ramsey 2005). Intersections of the roads are then nodes and connecting streets are edges of a graph. Each edge is then assigned a cost to be traveled, which in the simplest case is the length of the street segment. Choosing the closest nodes of the network to the random sampled location from the strata, or alternatively directly sampling nodes, an optimized path along the network visiting all the chosen nodes can be calculated using the Dijkstra algorithm (Dijkstra 1959) as implemented in pgRouting (pgRouting Contributors 2013)(see also Wieland et al. 2012; Pittore et al. 2015). If indicated in the street network data, pedestrian streets or one way connections can be considered. For efficiency, one can try to avoid taking the same road twice by increasing the cost of an edge connection if it was used previously. In case limited resources are available, the previous steps can be used in an iterative procedure. Providing a target length of the route, one can start from a minimum sample, estimate the optimal route and keep adding sample points until the target length has been reached. This way, one obtains an optimized path through the town, covering as much as possible all of the different built-up strata types.

B.4. Taxonomy and rapid visual screening

The omni-directional imagery obtained following such an optimized path (see Figure B.4) can be analyzed for visible buildings. This procedure is referred to as remote rapid visual screening (RRVS). Since the images are geocoded, the engineer screening a building is most likely able to determine its location and geocode building information. If footprints of the buildings are available (e.g., from OSM), the building information can also be directly attributed to these. Since the mapped path is based on the initially chosen sampling design, one is able to obtain a rep-

representative sample of the built environment, allowing the derivation of the spatial distribution and frequency of the different building types. The assignment of such building types is usually based on a building taxonomy, which sets out to describe and classify buildings in a standardized and unambiguous way. If a faceted taxonomy such as the GEM building taxonomy (Brzev et al. 2013) is employed a rich description of the building stock can be achieved. Fundamentally, this relocates the decision to assign a building type in the field to a later stage when the collected data is analyzed and opens a multitude of improvements regarding the resulting exposure model.

The intended application of such taxonomies is within the framework of rapid visual screenings so as to quickly assess large numbers of structures. This means all features should be described through a visual analysis, without requiring any measurements. Such data is not sufficient to model the response of a structure to loading, which requires detailed information on the geometry, used materials and material properties, usually only obtainable through thorough analysis and measurements of the structure itself. Such rapid screening methods should thus ideally be complemented by the detailed analysis of individual structures. The information retrieved this way can then be used to constrain the larger scale models. Nevertheless, individual building accuracy cannot be achieved this way and careful consideration of the resulting uncertainties is fundamental for statistically accurate risk estimates.

A simple way to perform rapid visual screening is through the analysis of mobile mapping imagery. Using such a system opens the possibility to perform the screening from remote (Remote Rapid Visual Screening, RRVS).

B.5. RRVS web platform

A simple web platform has been developed that allows the centralized collection of screening data via the internet. Figure B.4 shows a screenshot of the simple interface of the tool. It consists of three panels. The top right panel, shows an interactive map with a satellite view of the town (Mapbox, OSM) and a building (polygons, OSM), as well as an image location layer (points). The top left of the interface is reserved for an interactive panoramic image viewer which shows the image at the location selected from the map. The lower part of the interface contains the data entry form, comprised of several tabs for different building attribute categories and offers drop down menus and text entry fields for the different building attributes according to the GEM Building Taxonomy v 2.0 (Brzev et al. 2013). Given credentials, i.e., a user and task id, a user can log in via an internet connection, using only a simple web browser. The users start a survey by logging into the system. They can then select a building for screening either via a table of buildings within the task or by clicking on a building footprint in the map. Doing so populates the interface with all information stored about the building in the database (by default, all attributes are set to unknown) and the user can start modifying them. Selecting a building automatically loads the closest available image into the panoramic viewer and tries to orient the view towards the building. The user can also select other images by clicking on image locations on the map. Finally, after clicking on an update button, the provided information is stored in a database and the building's status changes from UNMODIFIED to MODIFIED or COMPLETED (the latter only if ticked in the interface), with the building footprint changing color in the map from red to yellow to green, respectively. Each building can be revisited at any point. In some cases, it is useful to compare the available images with images that might be available on Google Street View, thus a link is embedded in the panoramic viewer which links to the closest Google Street View location (if available).

A task is usually comprised of 100 buildings and is generated by randomly sampling all buildings along the mobile mapping path, i.e., where images are located close by. This way, every completed task complements previously completed tasks and increases the knowledge about a town globally. The system can be used by several users at the same time. Based on limited experience with the developed system, an analyst, depending on skill and personal desire for detail, can survey between 10 and 20 buildings per hour. Of course, this excludes any preparation and processing time to collect the images and prepare the survey tasks, which is probably higher than for more traditional approaches. On the other hand since all data is collected in a standardized and centralized way, post-processing, analysis and sharing of the collected information is relatively simple.

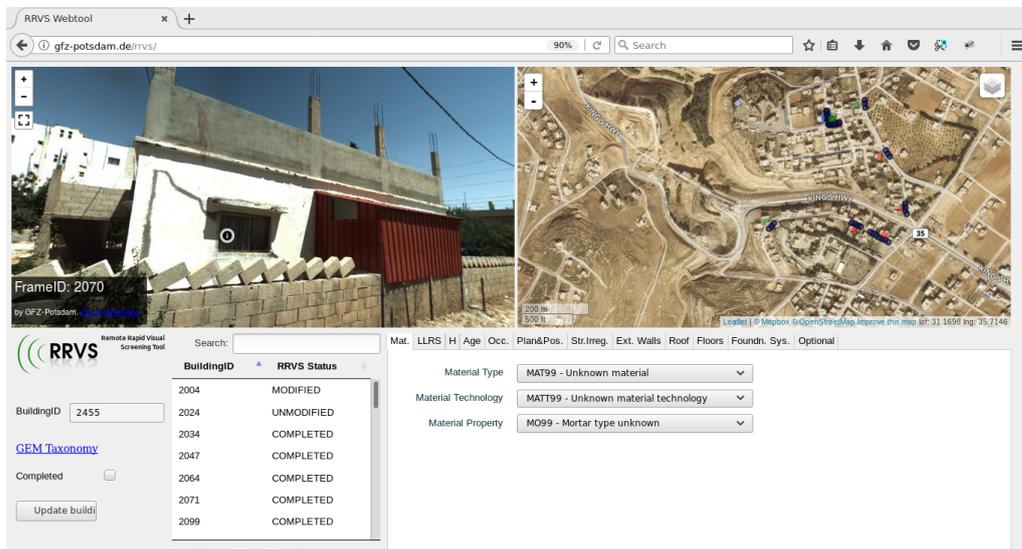


Figure B.4.: Screenshot of the RRVS web tool, composed of an overview map (top right) showing a set of buildings (polygons) and available panoramic images nearby (points), a viewer for these images (top left) and the data entry form for attributes based on the GEM Building taxonomy.

B.6. GEM Taxonomy Codes

Table B.2.: Short codes used for attributes within this study (please refer to Brzev et al. 2013, for a complete list of GEM Building Taxonomy codes, rows with * are not part of the taxonomy).

Code	Description
AGR1	Produce storage
AGR2	Animal shelter
AGR3	Agricultural processing
AGR99	Agriculture, unknown type
ASS1	Religious gathering
ASS2	Arena
ASS3	Cinema or concert hall
ASS4	Other gatherings
ASS99	Assembly, unknown type
BPD	Detached building
BT0*	No basement
BTS*	Shallow basement (1 storey below ground)
BTD*	Deep basement (2+ stories below ground)
C99	Concrete, unknown reinforcement
CHV	Change in vertical structure (includes large overhangs)
CIP	Cast-in-place concrete
COM1	Retail trade
COM10	Airport
COM11	Recreation and leisure
COM2	Wholesale trade and storage (warehouse)
COM3	Offices, professional/technical services
COM4	Hospital/medical clinic
COM5	Entertainment
COM6	Public building
COM7	Covered parking garage
COM8	Bus station
COM9	Railway station
COM99	Commercial and public, unknown type
CR	Concrete, reinforced
CU	Concrete, unreinforced
DNO	Non-ductile
E99	Earth, unknown reinforcement
EDU1	Pre-school facility
EDU2	School

Continued on the next page

Table B.2 – continued from previous page

Code	Description
EDU3	College/university offices and/or classrooms
EDU4	College/university research facilities and/or labs
EDU99	Education, unknown type
ER	Earth, reinforced
EU	Earth, unreinforced
EWC	Concrete exterior walls
EWMA	Masonry exterior walls
FC	Concrete floor
FC2	Cast-in-place beam-supported reinforced concrete floor
FME2	Metal floor beams supporting precast concrete slabs
FOSN	Shallow foundation, no lateral capacity
GOV1	Government, general services
GOV2	Government, emergency services
GOV99	Government, unknown type
HTHR*	Height type high rise (10+ stories)
HTLR*	Height type low rise (1-2 stories)
HTMR*	Height type mid rise (3-9 stories)
IND1	Heavy industrial
IND2	Light industrial
IND99	Industrial, unknown type
IRIR	Irregular structure
IRPP	Plan irregularity - primary
IRVP	Vertical irregularity - primary
LFBR	Braced frame
LFINF	Infilled frame
LWAL	Wall
M99	Masonry, unknown reinforcement
MATO	Other material
MCF	Masonry, confined
ME	Metal (except steel)
MIX	Mixed use
MIX1	Mostly residential and commercial
MIX2	Mostly commercial and residential
MIX3	Mostly commercial and industrial

Continued on the next page

Table B.2 – continued from previous page

Code	Description
MIX4	Mostly residential and industrial
MIX5	Mostly industrial and commercial
MIX6	Mostly industrial and residential
MIX99	Mixed, unknown type
MOC	Cement mortar
MOL	Lime mortar
MR	Masonry, reinforced
MUR	Masonry, unreinforced
PLFI	Irregular plan shape
PLFR	Rectangular, solid
Post93*	Built after the Oslo accord in 1993
Pre93*	Built before the Oslo accord in 1993
RC	Concrete roof
RC2	Cast-in-place beam-supported reinforced concrete roof
RES1	Single dwelling
RES2	Multi-unit, unknown type
RES2A	2 Units (duplex)
RES2B	3-4 Units
RES2C	5-9 Units
RES2D	10-19 Units
RES2E	20-49 Units
RES2F	50+ Units
RES3	Temporary lodging
RES4	Institutional housing
RES5	Mobile home
RES6	Informal housing
RES99	Residential, unknown type
RME2	Metal roof beams supporting precast concrete slabs
RMN	Concrete roof without additional covering
RSH1	Flat
RWC99	Roof-wall diaphragm connection unknown
S	Steel
SRC	Concrete, composite with steel section
ST99	Stone, unknown technology
TOR	Torsion eccentricity
W	Wood

Continued on the next page

Table B.2 – continued from previous page

Code	Description
YAPP	Approximate date of construction or retrofit

B.7. Clustering exposure data to reduce complexity

An unsupervised clustering approach is applied to a set of buildings and their observed attributes in order to find homogeneous groups of buildings among them. Considering n buildings with p attributes, the building observations can be represented as a n -by- p matrix with each row corresponding to a building and each column corresponding to an attribute (Kaufman and Rousseeuw 2005, p. 4). In order to cluster such a data set, one requires a proximity measure for all pairs of the n buildings (Kaufman and Rousseeuw 2005, p. 4), i.e., a n by n matrix of similarities or dissimilarities. There are numerous proximity measures available but they clearly depend on the type of data, i.e., if a variable is continuous, ordinal or categorical. While finding the proximity of two values of a continuous variable is straight forward, e.g., via the Euclidean distance, this is more difficult for ordinal and seems hardly feasible for categorical variables. Most of the attributes of building data defined with the GEM Building Taxonomy are categorical. Different measures to define the similarity between individual categorical building attributes can be explored: a simple method based on a binary measure and a continuous measure derived using natural language applied to a very short textual description of the attribute values as well as on longer articles.

Usually, for categorical variables, binary measures are used, i.e., two values are similar only in case they are the same. Translating this to buildings defined with the GEM taxonomy, the simplest approach to consider the similarity $s_{ij;G}$ between a single pair of attribute values for building i and j , based on the two individual attribute values k_i and k_j of an attribute k , is (Gower 1971):

$$\sigma_{i,j;G} = \begin{cases} 0 & , \text{if } k_i \neq k_j \\ \sigma_G & , \sigma_G \in (0, 1) \text{ if } k_i \sim k_j \\ 1 & , \text{if } k_i = k_j \end{cases} \quad (\text{B.1})$$

Since this describes the similarity of the two buildings only with respect to the variable k , for the case that all p attributes are comparable, the overall similarity between building i and j , $s_{ij;G}$, is defined as (Gower 1971):

$$s_{ij;G} = 1/p \sum_{k=1}^p s_{ijk;G} \quad (\text{B.2})$$

which when closer to one, the more similar two buildings are, and when closer to zero otherwise.

Of course, one might argue that some of the categorical variables in the GEM building taxonomy are more similar than others, e.g., one can consider the material type, CR (Reinforced concrete) is more similar to $C99$ (Concrete with unknown reinforcement) or even CU (Unreinforced concrete) than URM (Unreinforced masonry). Employing the Gower similarity in this case, if the first building's attribute is of the first type and the second's is of one of the latter, the attribute pair will in all cases have a similarity value $s_{ijk} = 0$ for this variable. Now, a possibility to relax this strict similarity definition could be to manually assign similarities. While this

might still be feasible for small sets of attribute values, this problem quickly reaches an impractical size, considering that there are about 300 different attribute levels in the taxonomy. Thus, an alternative way to obtain a more flexible and continuous similarity measure is proposed in the following.

This alternative method is available to us as the attributes are collected using pre-defined categorical values from the GEM Building Taxonomy, meaning that one knows beforehand most attribute values a building can have. Consequently, one can exploit additional information on each attribute. For each attribute defined in the taxonomy, a short description is provided (Brzev et al. 2013). In addition GEM maintains an online glossary with a short article with more detailed descriptions. The fundamental idea of the proposal here is that using standard natural language processing techniques, one can find the textual similarity between two texts describing individual attributes. If the texts are defining the attribute sufficiently precise one should be able to use this textual similarity as an approximation for the attribute similarity.

In order to implement this idea, both textual descriptions of the attributes are collected. While the short descriptions can be obtained from Brzev et al. (2013), a small script which automatically searches the GEM Glossary website for all attribute values and gathers for each the individual glossary article text is used. This yields two sets of texts for each of the attributes. In order to estimate the text similarity, both text sets have to be vectorized using the bag of words approach (see e.g., Joachims 2002, 13f). This corresponds to collecting the unique set of all terms present in all the articles and then for each article binary encoding presence of each of the terms with a value of 1, if present, and 0 otherwise. The resulting vector has the same length for each attribute text with mainly zero values. Removing all punctuation, numbers, 172 common words in English (excluding "above" and "below" from the "stopwords" list implemented in Feinerer et al. 2008) and lemmatizing conjugated words yields relatively clean vectors for each text.

For each of the text sources, a so-called term document matrix (TDM) is composed where rows are the vectorized texts for the 300 attributes. The longest text obtained from the short descriptions has 17 terms and thus the TDM obtained from it is rather compact with only 540 unique terms corresponding to its columns. As of August 2017, longer texts from the GEM Glossary (number of characters 26-4'913; mean: 468, median: 253, 95%: < 1'500s) result in a TDM with 1'818 different terms.

To obtain a similarity measure s for any two vectorized textual descriptions (\vec{v}_i, \vec{v}_j), the cosine similarity is calculated from each combination of the vectorized attribute value descriptions using the definition of the dot product (see e.g., Dale et al. 2000, p. 471):

$$\vec{v}_i \cdot \vec{v}_j = \|\vec{v}_i\| \|\vec{v}_j\| \cos(\theta) \quad (\text{B.3})$$

which after rearranging yields:

$$\cos(\theta) = \frac{\vec{v}_i \cdot \vec{v}_j}{\|\vec{v}_i\| \|\vec{v}_j\|} = s(v_i, v_j). \quad (\text{B.4})$$

Please note, that negative term frequencies are impossible as the vectors are binary,

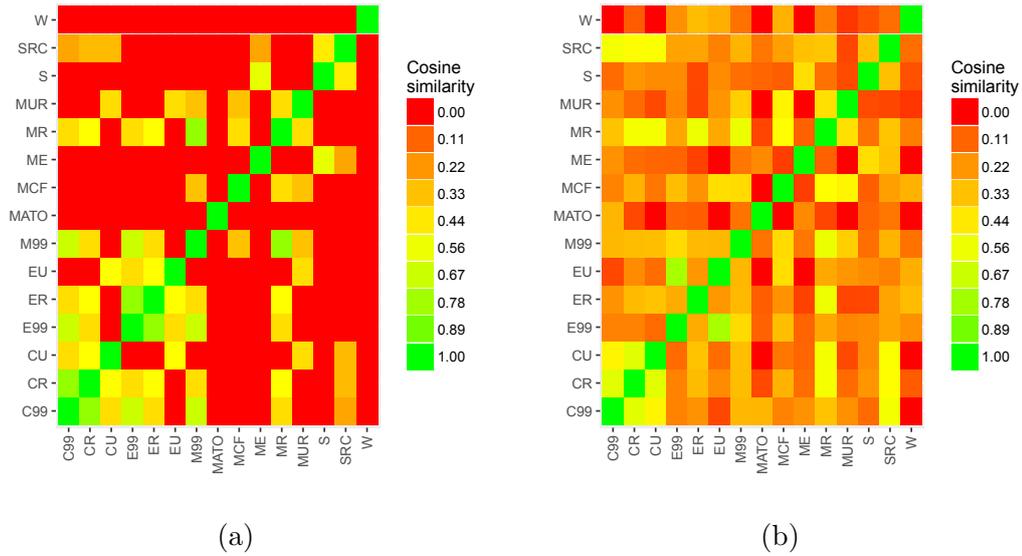


Figure B.5.: Subset of the resulting similarity matrices for the attribute 'Material Type'. (a) the cosine similarity approach using the short textual descriptions as in Brzev et al. (2013), and (b) attribute articles from the GEM Glossary. For a list of the attribute value codes used in this thesis, see Table B.2

equation B.4 will thus only return values between 0 and 1 and can be used directly as a similarity measure for \vec{v}_i and \vec{v}_j , where 0 indicates dissimilarity and 1 indicates full similarity.

The hypothesis is that the cosine similarity $s(v_i, v_j)$ of the textual description of attribute values i and j can be used as an approximation for the attribute similarity $s_{i,j;C} \sim s(v_i, v_j)$. For each value of an attribute, the textual cosine similarity to a neighboring attribute value is assigned as the similarity between attribute values.

While the binary Gower similarity as described in the beginning of the section can be seen as a unity matrix, Figure B.5 shows the continuous distributions derived with the cosine similarity. For readability reasons, only a subset of the matrices, for the attribute *Material type*, is shown (please refer to Table B.2 for a short description of each attribute code used in the figure and throughout the study). Figure B.5a shows the similarity measure obtained from the short textual description (Brzev et al. 2013) and B.5b shows the one derived using the GEM Glossary.

Judging from the similarities based on the short description as shown in B.5a, one can see that in some cases, the textual similarity seems sufficient to characterize attribute similarity, e.g., metal (*ME*) is considered similar to steel (*S*) and to some extent to steel with reinforced concrete (*SRC*). In other cases, it is clear that the short textual description is not sufficient to reliably determine all similarities. An example is concrete with unknown reinforcement (*C99*) which due to the terms "unknown" and "reinforcement" is considered more similar to earthen structures with unknown reinforcement (*E99*) and masonry with unknown reinforcement (*M99*) than to unreinforced concrete (*CU*). Using the longer article texts of the GEM on-

line glossary (Figure B.5b), reduces these effects but it also introduces some scatter. It assigns more similarity between confined masonry (*MCF*) and reinforced masonry (*MR*) than for its unreinforced counterpart (*MUR*). In addition, it can be observed that all concrete material types are almost equally similar to steel with reinforced concrete structures, where only slightly less similarity is given to the reinforced one. Reinforced masonry is considered almost equally similar to reinforced/unreinforced concrete, confined masonry and masonry with unknown reinforcement (*M99*).

Table B.3 shows the 5 most similar attribute values for different attributes. Some values are well defined, e.g., for material technology other steel (*SO*) and other metals (*MEO*) and for lateral load resisting system infilled frames (*LFINF*) and moment frames (*LFM*) are reasonably considered to be most similar. On the other hand, for material type, reinforced concrete is wrongly considered equally similar to unreinforced concrete as to concrete with unknown reinforcement and presence of soft storey (*SOS*) is considered similar to torsional eccentricity (*TOR*).

Likely sharper definitions or the usage of semantic algorithms could yield better results. Despite these clear deficiencies, one can use the derived similarities to cluster a building data set collected according to the GEM Building Taxonomy. In order to do so, the similarity measures must be turned into a dissimilarity/distance, i.e., subtracting each value from 1. In addition, one might consider some attribute k more relevant from a structural point of view than others. The simplest way to account for this is the application of a weighting scheme, by multiplying the similarity with a scalar attribute weight w_k to obtain a weighted similarity:

$$s_{ijk}^* = w_k s_{ijk}. \quad (\text{B.5})$$

After calculating the distances for all buildings in a dataset according to any of the dissimilarity measure introduced above (Gower or cosine similarity), one can apply a clustering approach such as the k-mean clustering previously used for the selection of ground motion models.

Table B.3.: Most similar attribute levels (similarity in braces) for several attributes using the GEM Glossary articles.

Attribute	Attribute pairs and cosine similarity
MAT_TYPE	E99~EU (0.77)
	CR~CU (0.6)
	C99~CR (0.6)
	M99~MR (0.57)
	ER~MR (0.56)
MAT_TECH	MEO~SO (0.85)
	ETO~SO (0.8)
	ETO~MEO (0.78)
	ST99~STRUB (0.71)
	CL99~CLBRS (0.66)
MAT_PROP	MOL~MOM (0.64)
	MOC~MOCL (0.63)
	MOC~MOM (0.58)
	BOL~WEL (0.53)
	MOCL~MOL (0.52)
LLRS	LFINF~LFM (0.78)
	LFLS~LFLSINF (0.62)
	LDUAL~LWAL (0.56)
	LDUAL~LH (0.55)
	LFBR~LFM (0.53)
OCCUPY	EDU~RES (0.57)
	EDU~GOV (0.47)
	IND~RES (0.45)
	AGR~RES (0.41)
	COM~EDU (0.4)
OCCUPY_DT	RES2B~RES2C (1)
	RES2B~RES2D (1)
	RES2C~RES2D (1)
	RES2B~RES2E (1)
	RES2C~RES2E (1)
STR_IRREG_TYPE	SOS~TOR (0.58)
	IRHO~IRVO (0.48)
	REC~TOR (0.47)
	POP~SOS (0.4)
	POP~TOR (0.4)

C. Fragility related

C.1. Building models for Nablus

For the purpose of this study, the approximately 5% of buildings in Nablus, which are not made of reinforced concrete are excluded and the many structural irregularities, as well as the 5% of the buildings with seismic design are not considered. Building parameters were collected from the data presented here and additional sources (Al-Dabbeek and Al-Jawhari 2002; Grigoratos et al. 2016) in order to obtain structural models for the building types found in Nablus. These model should allow to derive simplified capacity and state-dependent fragility models.

Determining the structural parameters for these types, plan dimensions are obtained running a principal components analysis on their footprints (taken from OpenStreetMap OpenStreetMap contributors 2017) and imposing rectangular shape (65% of the buildings according to the data set employed here). On other structural parameters the most recent source of information is Grigoratos et al. (2016), who collected expert judgment of local engineers. Using this source and Al-Dabbeek and Al-Jawhari (2002), building parameters as shown in Table 5.4 in chapter 3.3 are chosen.

Only recently in 2016, has an official building code been introduced in Palestine. Thus, it is unclear which design procedure was followed for the buildings in the dataset. Therefore, for demonstration purposes of the proposed framework only Eurocode 2 (EC-2; European Committee for Standardization 2004a) based design is considered, i.e., only designed for gravitational loads. From the parameters and a simplified design procedure moment frame building models following EC-2 are constructed, assuming that no seismic design and damage hierarchy have been followed.

Using a design load resulting only from the slab mass, and $2t/m^2$ live loads (Eurocode-1 residential category A, European Committee for Standardization 2001), load levels on beams and supporting columns are determined. Each beam is designed with an initial depth of 60 *cm* and a width of 25 *cm*, reported as typical for Palestine in Al-Dabbeek and Al-Jawhari (2002). According to the beam design procedure as described in Lee et al. (2000, p. 37), the bending moment of the beam and required tensional reinforcement in the beam are determined. For simplicity it has been ensured that no compressional reinforcement is required by increasing the beam dimensions in 10 and 5 *cm* steps, respectively, if necessary to balance the moments. Imposing square design, column dimensions are chosen at the slenderness limit for

stocky columns (Bond et al. 2006, p. 36), conservatively choosing story height as effective column length. The longitudinal column reinforcement is then derived following EC-2 based design charts in Lee et al. (2000), balancing the moments at the column-beam joint. The standard rebar diameter is chosen as 16 *mm* and only multiples of 4 are allowed as total number of rebars. To adhere for the latter, the reinforcement ratio may exceed the design requirements, but was set to not exceed these by more than 50%. Constraining the column reinforcement area according to the requirements of EC-2 (9.5.2), a maximum and minimum are set to $A_{s,max} = 0.04A_g$ and $A_{s,min} = 0.002A_g$. In case the configuration is found insufficient to these criteria, the diameter of the rebars and/or the column's dimensions are adjusted until a fitting combination is found, where the rebar diameter is changed in 4 *mm* steps (minimum diameter 8 *mm*, EC-2 9.5.2-1) and column dimensions in 5 *cm* steps. An algorithm was developed to design the building models from the provided parameter distributions by sampling the distributions, using Latin Hyper Cube sampling. For simplicity, only a single median or fixed sample of the parameter distributions as indicated in Table 5.4 was used to build the models for Nablus. This yields 48 different building models. Combining the three different slab types, the 48 configurations are grouped into 16 subtypes composed of 4 age classes (before 1980, 1980s, 1990s and since 2000) and 4 building heights to represent low rise (2 floors), lower (4) and upper mid-rise (7), and high rise (10). The only parameter variation is thus the slab type.

C.2. Eurocode 8 building capacity for moment frame buildings

In order to assess a building's fragility with regard to seismic loading, the lateral capacity of the structure has to be determined. A simplified approach to approximate the lateral capacity of moment frame buildings is described here, considering only the columns of the structure. This is a worst case assumption of weak columns and strong beams, contrary to modern building design. It assumes that no damage hierarchy was considered in the building design, meaning that structural failure forms in the columns rather than the beams. Contributions of walls to mass and stiffness, as well as the self-weight of columns, are disregarded for the analysis. For each individual column, the capacity is analyzed following Part 3 of Eurocode 8 (EC-8; European Committee for Standardization 2004b), once in each direction perpendicular direction of loading, considering no safety factors in the calculations. Regarding the failure mechanism brittle as well as ductile column failure is considered.

A column's geometry is defined by its height H , width b and depth h_c of its cross section. The longitudinal steel reinforcement within the column is described by its location within the column cross section with respect to h_c and b and its diameter d_s (Figure C.1a). The distance between the center of the outermost steel reinforcement and the rim of the column is denoted as c . As shown in figure C.1b, L_s is the lap splice length of the longitudinal reinforcement and the transverse reinforcement is defined via its diameter d_w and the stirrup spacing s . The effective depth of the

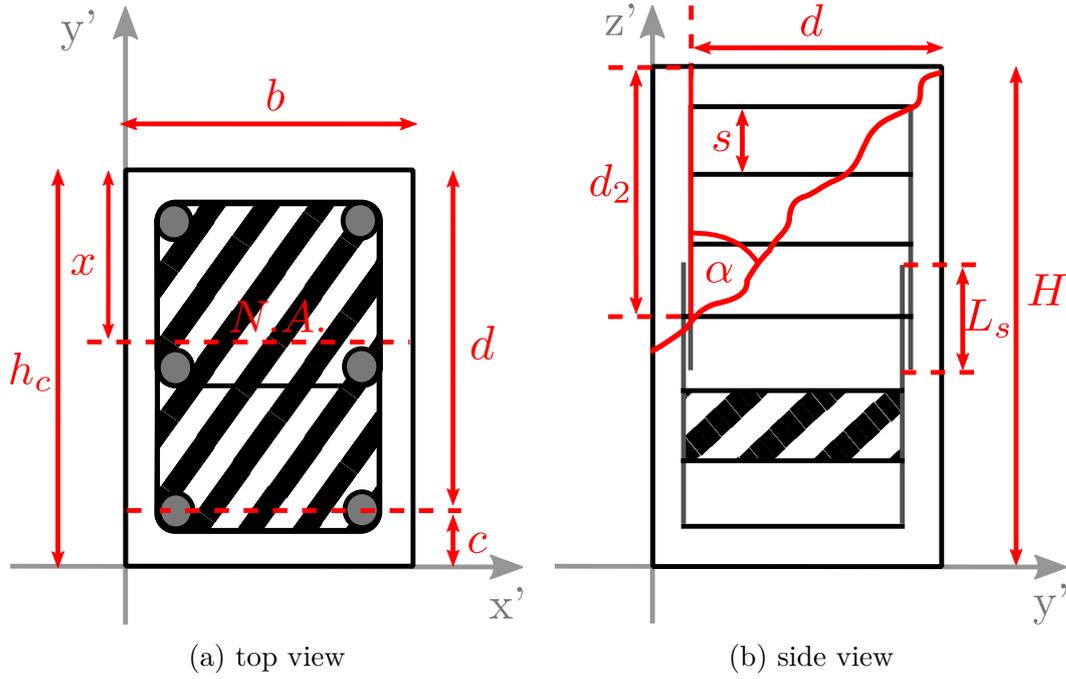


Figure C.1.: Cross sections of a column with height H , width b , and depth h_c . Longitudinal reinforcement (gray), and the area (striped) confined by the stirrups (black). Figure C.1a shows a cross section in the $x'y'$ -plane with the compression zone depth x to the neutral axis ($N.A.$), the thickness of the concrete cover c , and the effective depth d . Figure C.1b shows a column cross section in $y'z'$ -plane with the stirrups (vertical black lines), the longitudinal reinforcement (horizontal black lines), and the area (striped) confined by two stirrups with spacing s . Lap splicing is indicated with the splice length L_s . The red line at an angle α indicates a potential crack due to shear, where $d_2 = d$ in case $\alpha = 45^\circ$

cross section is $d = h_c - c$.

First step in the analysis is the determination of the compression zone depth x . For a column cross section loaded with N at $0.5h_c$ (figure C.1a), x is determined iterative as the one yielding a stress equilibrium between the concrete under compression and steel under tension, neglecting concrete tensile strength. In order to do so, the concrete cross section is split in layers with height $dy = 10^{\lfloor \log_{10}(c) \rfloor}$ (dotted rectangle in Figure C.2a)

and it is assumed, that the tensile reinforcement that is the furthest away from the neutral axis (red colored), i.e., has the longest lever arm $l_i(x) = y_i - (h_c - x) = l_{max}$ (depending on the imposed x) has reached its yielding strain $\epsilon_{s,y}$. The latter is estimated from the stress level at which the steel reaches yielding $f_{s,y}$ and Hooke's law for its elasticity modulus E_s (Beer et al. 2011, p. 62):

$$\epsilon_{s,y} = \frac{f_{s,y}}{E_s} \quad (C.1)$$

From the yielding strain and the maximum lever arm l_{max} , the curvature of the cross

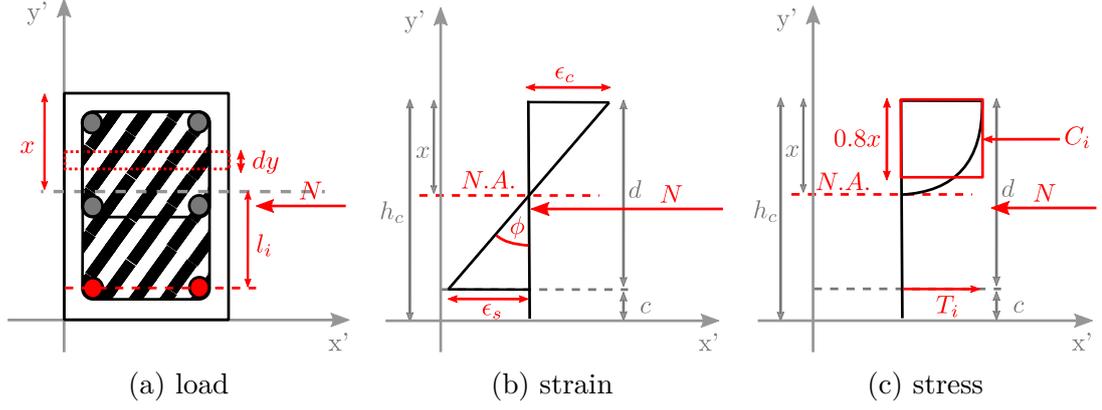


Figure C.2.: Loaded cross section, strain diagram and stresses. To iteratively determine the compression zone depth, the concrete is split in layers with height dy (Figure C.2a). Curvature Φ (Figure C.2b) at yielding from the assumption that the tensile reinforcement with maximum distance l_i from the neutral axis reached a strain ϵ_s equal to its yielding strain. From this the tensile stress T_i (Figure C.2c) is calculated for each layer of reinforcement, considering that the parabolic compression stress distribution can be approximated with a rectangular distribution with depth $0.8x$. Calculating the strain using the curvature Φ for each concrete layer, and from these the compression stresses C_i , considering a strength and strain change in confined sections (striped area; see also Figure C.1b). The desired compression zone depth x is found, if the tensile and compressive forces calculated from the C_i s, T_i s and affected areas yield the applied load N .

section under the load is calculated (Beer et al. 2011, p. 232):

$$\Phi_y = \frac{\epsilon_{s,y}}{l_{max}} \quad (C.2)$$

Using the curvature, for each layer of concrete the strain at the center of the layer is:

$$\epsilon_j = \frac{\Phi_y}{l_j} \quad (C.3)$$

with lever arm l_j . Using the Hognestad (1951) concrete stress model, the layer stress is:

$$\sigma_j = \begin{cases} f_c & ; \forall \epsilon_j < \epsilon_{c,y} \\ f_c \left(\frac{2\epsilon_j}{\epsilon_{c,y}} - \left(\frac{\epsilon_j}{\epsilon_{c,y}} \right)^2 \right) & ; otherwise \end{cases} \quad (C.4)$$

where f_c and $\epsilon_{c,y} = 0.002$ (EC-2 Table 3.1) are concrete compressive strength and yielding strain. For concrete strength, confined parts (within the stirrup; striped area in figure C.2a) are distinguished from the unconfined part by considering a change in f_c to f_{cc} and $\epsilon_{c,y}$ to $\epsilon_{cc,y}$ as defined in EC-8 Part 3 (A.6):

$$f_{cc} = f_c \left[1 + 3.7 \left(\frac{\alpha_{cf} \rho_{sx} f_{w,y}}{f_c} \right)^{0.86} \right] \quad (C.5)$$

where $f_{w,y}$ is the yield strength of the transverse reinforcement, taken equal to $f_{s,y}$, $\rho_{sx} = A_{sx}/(b_w s)$ is the ratio of transverse steel parallel to the loading direction, with A_{sx} the cross section areas of the stirrup legs parallel to the loading, and $b_w s$ is the area parallel to the loading confined by two stirrup layers (see striped area in figure C.1b), where $b_w = b - 2c$ (analogous h_w). The factor α_{cf} accounts for the fact that the effectively confined volume in between the stirrups of rectangular columns is not a cuboid, but its edges are parabolas bent inwards (EC-8 Part 3 A.2):

$$\alpha_{cf} = \left(1 - \frac{s}{2b_w}\right) \left(1 - \frac{s}{2h_w}\right) \left(1 - \frac{\sum b_i^2}{6h_w b_w}\right) \quad (C.6)$$

From f_{cc} the yielding strain for the confined section is (EC-8 Part 3 A.7):

$$\epsilon_{cc,y} = \epsilon_{c,y} \left[1 + 5 \left(\frac{f_{cc}}{f_c} - 1\right)\right] \quad (C.7)$$

With the optimized x , the nominal moment capacity at yielding M_y is estimated, considering that the stress distribution is parabolic and can be approximated by a rectangular distribution with compression zone depth $0.8x$ (EC-2 3.19). Thus, the load N applied at the center of gravity, i.e., $0.5h_c$ creates a yielding moment (Pantazopoulou and Syntzirma 2010, p. 362):

$$M_y = \sum_i A_{i;s,t} f_{s,y} l_i + N(0.5h_c - 0.4x), \quad (C.8)$$

with $A_{i;s,t}$ the area of tensile reinforcement with yield strength $f_{s,y}$ at each layer of reinforcement at level l_i , with a lever arm from the center of the steel to the compression zone center at $0.4x$ from the top. The second term is the contribution of the concrete with a lever arm from the center of gravity to the compression zone center, i.e., $0.5h_c - 0.4x$.

Following EC-8, ductile column chord rotation (drift) at the onset of plastic deformation, i.e., yielding is defined as (EC-8 Part 3 A.10b):

$$\Theta_y = \Theta_{y,flex} + \Theta_{y,hear} + \Theta_{y,slip}, \quad (C.9)$$

where the three terms represent flexural, shear and slip contribution. The first term, i.e., flexural rotation is calculated as (EC-8 Part 3 A.10b):

$$\Theta_{y,flex} = \frac{1}{3} \Phi_y (L_v + a_v x) \quad (C.10)$$

with Φ_y , the yield curvature as estimated during the compression zone analysis.

The shear span L_v in equation C.10 is taken as $L_v = 0.5H$ and the variable a_v is a switch if shear cracking is expected before flexural yielding, i.e., $a_v = 1$ if $M_y > L_v V_c$, where V_c is the concrete shear resistance and $a_v = 0$, otherwise. V_c is estimated according to (EC-8 Part 3 A.12):

$$V_c = \frac{h_c - x}{2L_s} \min(N; 0.55A_c f_c) + (1 - 0.05 \min(5; \mu_{\Delta}^{pl})) \cdot \left[0.16 \max(0.5; 100\rho_{tot}) \left(1 - 0.16 \min\left(5; \frac{L_v}{h_c}\right)\right) \sqrt{f_c} (b - 2c) d 10^3 + V_w\right], \quad (C.11)$$

where $A_c = (b - 2c)d$ is the effective shear cross section, μ_{Δ}^{pl} is the ductility ratio, defined as the plastic chord rotation normalized by the yielding chord rotation, i.e., $\mu_{\Delta}^{pl} = 1$ for yielding. V_w is the shear resistance contribution of the transverse reinforcement, calculated as the shear resistance of all stirrups crossed by a shear crack with angle $\alpha = 45^\circ$ (corresponding to the maximum shear stress plain, see figure C.1b, where $d_2 = d$ with $\alpha = 45^\circ$) (EC-8 Part 3 A.13):

$$V_w = \frac{d}{s} A_w f_{w,y} \quad (C.12)$$

with A_w the cross section area of a stirrup calculated as:

$$A_w = \frac{\pi}{4} d_{ws}^2. \quad (C.13)$$

The second term $\Theta_{y,shear}$ in equation C.9 is the shear contribution to ductile failure, i.e. (EC-8 Part 3 A.10b):

$$\Theta_{y,shear} = 0.0013 \left(1 + 1.5 \frac{h_c}{L_v} \right), \quad (C.14)$$

the third term is the rotation due to slip at the anchorage of the column, due to pull out of the reinforcement and calculated as (EC-8 Part 3 A.10b):

$$\Theta_{y,slip} = 0.13 \Phi_y \frac{d_s f_{y,s}}{\sqrt{f_c}}. \quad (C.15)$$

The flexural shear demand at yielding V_y is obtained from the shear span definition as ratio between yield moment and the base shear at yielding, i.e.:

$$V_y = M_y / L_v. \quad (C.16)$$

The ultimate chord rotation capacity without diagonal reinforcement can be calculated from EC-8 (A.1):

$$\Theta_u = 0.016 \cdot 0.3^v \left[\frac{\max(0.01; w') f_c}{\max(0.01; w) 10^6} \right]^{0.225} \left(\frac{L_v}{h_c} \right)^{0.35} 25^{(\alpha_{cf} \rho_{sx} f_{w,y} / f_c)} \quad (C.17)$$

where w' and w are the ratios of tensile and compressive longitudinal reinforcement with respect to the gross section of the column and $v = N/A_g$ is the load stress on the section.

Ultimate flexural shear demand V_u is obtained in analogy to equation C.16, where M_u is obtained using equation C.8 with $f_{s,y} = f_{s,u}$.

For the brittle failure mechanism maximum shear resistance is estimated using equation C.11, where, as the ductility ratio μ_{Δ}^{pl} increases, the shear strength decreases according to equation C.11 until complete strength loss, which the code sets at $\mu_{\Delta}^{pl} = 5$.

For each column, there are thus three possible configurations for its failure mode, the column is either experiencing a) brittle failure before flexural yielding, b) brittle failure after flexural yielding, or c) flexural failure.

For each story, the capacity is then determined considering columns as springs in parallel (Weggel et al. 2007). Thus, the shear capacity of a story at a given drift level is taken as the sum of the shear capacity at this drift level of each column in a story.

Using a simplified building capacity model, the base shear can then be obtained from the ground floor capacity curve. For buildings with more than one story, the roof displacement resulting from this base shear has to be determined. In the simplest case this means to determine the fundamental mode shape of the building and use its shape to translate base drift to roof drift. One way to approximate this shape is the Rayleigh method, using equivalent single degree of freedom systems (ESDOF) to represent the structure. The fundamental idea of an ESDOF representation of a multi degree of freedom (MDOF) structure is, that its 1-D displacement shape approximates the MDOF displacement shape and can be described as a simple harmonic oscillator with an eigenfrequency ω , obtained from a generalized mass m^* and generalized stiffness k^* . Recollecting, that the time (t) dependent equation of motion for a SDOF system loaded with $N(t)$ can be obtained from a force equilibrium (Clough and Penzien 1975, p. 18):

$$F_I(t) + F_D(t) + F_S(t) = N(t) \quad (\text{C.18})$$

with $F_I(t)$ being the inertial force, $F_D(t)$ the damping force of the spring, and F_S the restoring force. If no external load is considered, i.e., $N(t) = 0$, equation C.18 becomes a harmonic ordinary differential equation (Clough and Penzien 1975, p. 42):

$$m\ddot{u}(t) + c\dot{u}(t) + ku(t) = 0 \quad (\text{C.19})$$

with $u(t)$ the displacement, $\dot{u}(t)$ and $\ddot{u}(t)$ first and second derivative with respect to time, m its clumped mass, and k its stiffness. Using the general solution for this equation (Clough and Penzien 1975, p. 42):

$$\begin{aligned} u(t) &= Ae^{-i\omega t} \\ \dot{u}(t) &= -Ai\omega e^{-i\omega t} \\ \ddot{u}(t) &= A\omega^2 e^{-i\omega t} \end{aligned} \quad (\text{C.20})$$

and considering an undamped system with $c = 0$, the eigenfrequency of the SDOF is obtained as (Clough and Penzien 1975, p. 42):

$$\omega^2 = \frac{k}{m} \quad (\text{C.21})$$

Equation C.21 can be used to approximate the fundamental mode of a RC-Frame, if its displacement is represented employing an assumed displacement shape $\psi(x) = u(x, t)/Z(t)$, with $Z(t)$ the amplitude of displacement (Clough and Penzien 1975, p. 30) and using the structure's generalized mass m^* and stiffness k^* instead of m and k .

Considering only lumped masses m_i for each story, the generalized mass is the sum over m_i multiplied by each story's displacement shape (Clough and Penzien 1975, p. 31):

$$m^* = \sum_i^n m_i \psi_i^2. \quad (\text{C.22})$$

One can obtain the generalized stiffness from equation C.21 using the eigenfrequency ω approximated with the Rayleigh method. This method is based on the conservation of energy in a freely vibrating system without damping. Considering the RC-Frame and a first arbitrarily guessed displacement shape $\psi^{(0)}$ with a maximum displacement $Z_0^{(0)} = \max(Z(t)) = 1.0$ and $u_i = 1.0$ for all n stories, a first estimate for the maximum kinetic energy for the structure is gained as (Clough and Penzien 1975, p. 140):

$$\begin{aligned} T_{max}^{(0)} &= \frac{1}{2} \sum_i^n m_i (\dot{u}_i^{(0)})^2 \\ &= \frac{1}{2} \omega^2 Z_0^2 \sum_i^n m_i (\psi_i^{(0)})^2. \end{aligned} \quad (\text{C.23})$$

Similarly, the maximum potential energy, depending on the relative story to story displacement Δv_i , can be determined as (Clough and Penzien 1975, p. 140):

$$\begin{aligned} V_{max}^{(0)} &= \frac{1}{2} \sum_i^n k_i (\Delta u_i^{(0)})^2 \\ &= \frac{1}{2} Z_0^2 \sum_i^n k_i (\Delta \psi_i^{(0)})^2, \end{aligned} \quad (\text{C.24})$$

where k_i is the stiffness of a column, estimated according to Hooke's law as $k_i = V_y / (\Theta_y H)$. Assuming conservation of energy, i.e., equating C.24 and C.23, one obtains an approximate solution for the eigenfrequency ω . The accuracy of the estimated eigenfrequency depends on the assumed shape function $\psi(x)$. Any deviation from the real shape function requires external constraints to maintain the equilibrium. This increases the stiffness and strain energy, and consequently also the estimated frequency (Clough and Penzien 1975, p. 133). Considering this, the unrealistic choice of ψ gives an overestimate of ω , but the estimate can be improved using the maximum inertial forces generated by this first guessed deformation shape under the dead load of the structure (Clough and Penzien 1975, pp. 137,139):

$$p_i^{(0)} = \omega^2 m_i u_i^{(0)} \quad (\text{C.25})$$

and calculate improved deflections $u_i^{(1)}$ considering that the story shear $F_{u,i} = \sum_{j>=i}^n p_j^{(0)}$ is related to the deformation via the storey's stiffness as (Clough and Penzien 1975, p. 140):

$$\Delta u_i^{(1)} = \frac{F_{u,i}}{k_i}. \quad (\text{C.26})$$

Using the improved deflection shape, normalized to the value of the top story, and

rewritten as (Clough and Penzien 1975, p. 138):

$$\begin{aligned}
u_i^{(1)} &= \omega^2 \frac{u_i^{(1)}}{\omega^2} \\
&= \omega^2 \psi_i^{(1)} \frac{Z_0^{(1)}}{\omega^2} \\
&= \omega^2 \psi_i^{(1)} \bar{Z}_0^{(1)}
\end{aligned} \tag{C.27}$$

to compute the maximum potential energy produced by this loading with (Clough and Penzien 1975, p. 140):

$$\begin{aligned}
V_{max}^{(1)} &= \frac{1}{2} \sum_i^n p_i^{(0)} u_i^{(1)} \\
&= \frac{1}{2} \sum_i^n \omega^2 m_i u_i^{(0)} \omega^2 \psi_i^{(1)} \bar{Z}_0^{(1)} \\
&= \frac{\omega^4}{2} Z_0^{(0)} \bar{Z}_0^{(1)} \sum_i^n m_i \psi_i^{(0)} \psi_i^{(1)}
\end{aligned} \tag{C.28}$$

and the kinetic energy from (Clough and Penzien 1975, pp. 138,141):

$$\begin{aligned}
T_{max}^{(1)} &= \frac{1}{2} \sum_i^n m_i (\dot{u}_i^{(1)})^2 \\
&= \frac{\omega^2}{2} (Z_0^{(1)})^2 \sum_i^n m_i (\psi_i^{(1)})^2 \\
&= \frac{\omega^6}{2} (\bar{Z}_0^{(1)})^2 \sum_i^n m_i (\psi_i^{(1)})^2
\end{aligned} \tag{C.29}$$

and again equating both, yields a better approximation of ω (Clough and Penzien 1975, pp. 138,141):

$$\omega^2 = \frac{Z_0^{(0)} \sum_i^n m_i \psi_i^{(0)} \psi_i^{(1)}}{\bar{Z}_0^{(1)} \sum_i^n m_i (\psi_i^{(1)})^2} \tag{C.30}$$

With the mode shape derived this way, one can determine building capacity from the story capacities. Imposing displacements δ_{top} at the top of the structure, it can be determined which shear level this corresponds to at the base of the structure, i.e.,:

$$V_b(\delta_{top}) = V_1(\psi_1^{(1)} \delta_{top}) \tag{C.31}$$

where V_1 is the capacity of the ground floor in terms of displacement and $\psi_{1,1}$ is the improved mode shape of the ground floor. Converting this back to drift values via the building height H_b , one obtains the building capacity curve in the investigated loading direction.

Performing the analysis in both directions and choosing the weaker yields the building capacity curve. For the purpose of this thesis, a simple bilinear model is fit to this curve, expanding from the origin to the yielding point and from there until the

ultimate capacity point. In order to do so, first the ultimate drift Θ_u is defined as the point where the capacity drops to 85% of its maximum, then the yield drift Θ_y and capacity V_y point is found using a least squares optimization for the residuals between the resulting bilinear model and the determined capacity curve.

Capacity curves derived this way for Nablus are listed in Table 5.5.

C.3. Building response, deterioration and fragility

To determine building response a dynamic building response analysis using an incremental dynamic analysis (Vamvatsikos and Cornell 2002, IDA;) can be performed employing a linear ESDOF model, where the ESDOF parameters determined during the capacity analysis (Appendix C.2) are used. An established software solution to model dynamic response of structural elements is the free and open software OpenSees (McKenna 1997). It can be used to model an ESDOF's response using a columnBeamElement with the equivalent parameters of the building. In addition to equivalent mass, stiffness and eigenfrequency, this requires to define an equivalent height H^* , elasticity E^* , area A^* , and second moment of inertia I^* for the system. Using the determined mode shape, the equivalent height can be determined as (Priestley et al. 2007, p. 223):

$$H^* = \frac{\sum_i^n m_i \psi_i^{(1)} H_i}{\sum_1^n m_i \psi_i^{(1)}}. \quad (\text{C.32})$$

From the yielding concrete strength $f_{c,y}$ and strain $e_{c,y}$ (fixed to $e_c = 0.002$) the equivalent elasticity can be defined as:

$$E^* = \frac{f_{c,y}}{e_c}. \quad (\text{C.33})$$

Considering bending stiffness as:

$$K = F/\delta_d \quad (\text{C.34})$$

with applied Force F and resulting deflection δ_d , and the cantilever beam deflection for a beam with length a under a shear force F :

$$\delta_b = \frac{Fa^3}{3EI} \quad (\text{C.35})$$

one obtains for the equivalent inertia:

$$I^* = \frac{K^*(H^*)^3}{3E^*} \quad (\text{C.36})$$

At last, the equivalent area can be estimated from the equation for the second moment of inertia of a rectangular area:

$$I^* = \frac{ab^3}{12} \quad (\text{C.37})$$

assuming square dimensions, i.e., $a = b$.

Once a building experiences displacements beyond yielding, i.e., deforms plastically, the building response is non-linear and its dynamic properties change accordingly. This behavior can not be captured with a linear system. For the purpose of this thesis, thus the non-linear response is only approximated through a piecewise linear combination of single damage grade increments. Meaning, only transitions from

one damage state to the next are modeled within the IDA, using ESDOF models with stiffness and damping depending on the damage state. Hence, for each discrete damage state, that a building may have prior to an event (see equation 3.10) ESDOF parameters have to be estimated. Borrowing concepts from direct displacement based design (Priestley et al. 2007), the secant stiffness is determined for each limit state defined in the previous section using the respective drift value, and the capacity from the building capacity curve, i.e., (Priestley et al. 2007, p. 64):

$$K^*(S_{d,i}) = V_b(S_{d,i})/V_b(S_{d,i}) \quad (\text{C.38})$$

For the increase in damping an empirical relation for reinforced concrete frames by Priestley et al. (2007) based on the modified fat Takeda hysteresis rule is employed (Priestley et al. 2007, p. 86):

$$\xi^*(\mu) = \xi_{el} + 0.565 \frac{\mu - 1}{\mu\pi} \quad (\text{C.39})$$

with $\mu = S_{d,i>1}/d_y$ the ductility and $\xi_{el} = 0.05$ the elastic damping. Since the first limit state is still in the elastic range the stiffness and damping is the same as for the undamaged state.

As the whole framework used within this study is employing PGA, it is necessary to consider the variability in the ground motion to obtain the variability in the response of the ESDOF system. For fragility analysis the response needs to be estimated over the full range of ground motion levels of interest. The IDA simplifies this by linearly scaling amplitudes of few records to the full range. This of course is a simplification as in reality not only the amplitudes, but also other properties as, e.g., duration and frequency content are changing. In order to reduce the influence of this on the estimated fragilities, the selected records have to be collected carefully. Ideally, these are composed by a set of site-specific ground motions. Unfortunately, only few sites are a) monitored and b) monitored sufficiently long enough to have such a set of ground motions. Consequently records have to be selected, matching the site-conditions in terms of soil characteristics, but also the earthquake hazard at the site. In order to select records which are representative for the hazard conditions, a probabilistic seismic hazard model is required for the site. Disaggregating the PSHA curve for the site (Bazzurro and Cornell 1999) at, e.g., the 10% probability of exceedance within 50 years yields event parameters (distance to station, magnitude and deviation from the median ground motion), which are contributing most to this hazard level. These can be used to select records, whose station and causing event match these criteria. A simple method for the second requirement, i.e., to select records with the same site conditions is proposed in Eurocode-8. Site specific in terms of EC-8 means that the elastic response spectrum of the record matches the elastic response spectrum for the hazard level at the site of interest. The latter is obtained using a site-classification in terms of v_{s30} (see Table C.1) and requesting a zero period spectral response larger or equal than the a_g (equivalent to the *PGA* in $[g]$) obtained from the PSHA for a given design period. In addition, the code requires that no mean value of the 5% damped elastic spectrum of the selected records is less than 90% of the 5% damped elastic building response spectrum in the range between $0.2T_1$ and $2T_1$, where T_1 is the fundamental period of the oscillator.

Table C.1.: Eurocode-8 soil types and corresponding v_{s30} values (EC-8 Part 1 Table 3.1)

Ground type	Description of stratigraphic profile	$v_{s30}[m/s]$
A	Rock or other rock-like geological formation, including at least 5 m of weaker material at the surface	> 800
B	Deposits of very dense sand, gravel, or very stiff clay, at least several tens of meters in thickness characterized by a gradual increase of mechanical properties with depth.	360 – 800
C	Deep deposits of dense or medium-dense sand, gravel or stiff clay with thickness from several tens to many hundreds of meters.	180 – 360
D	Deposits of loose-to-medium cohesionless soil (with or without some soft cohesive layers), or of predominantly soft-to-firm cohesive soil.	< 180
E	A soil profile consisting of a surface alluvium layer with v_s values of type C or D and thickness varying between about 5 m and 20 m, underlain by stiffer material with $v_s > 800 m/s$.	
S_1	Deposits consisting, or containing a layer at least 10 m thick, of soft clays/silts with a high plasticity index ($PI > 40$) and high water content.	< 100 (only indicative)
S_2	Deposits of liquefiable soils, of sensitive clays, or any other soil profile not included in types A-E or S_1 .	

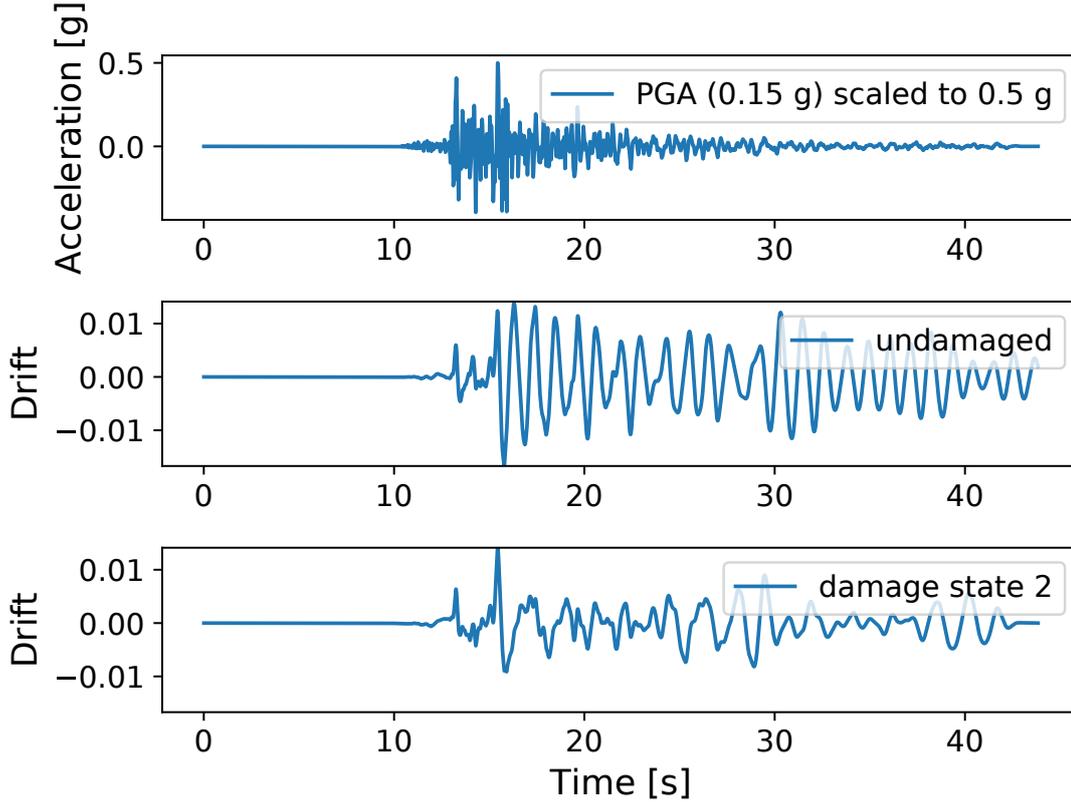


Figure C.3.: Ground acceleration recorded during an event at a station with soil-type B (EC-8) in Greece (HI.PAT1..HN2.D.19930714.123148.C.ACC.ASC) and response of an ESDOF system once approximating an intact building and once a extensively damaged version (ds_2 in equation 3.10 of a building). Please note that in the actual simulation not necessarily the same record is chosen for both systems as their eigen period might not match the EC-8 criteria.

In order to use the mean results of the analysis EC-8, at least seven time history response analyses have to be run.

Performing the IDA for each of the ESDOF models with the ground motion set, scaled to the different levels of PGA, one obtains drift time histories for each record. Figure C.3 shows a time-series of ground acceleration, recorded at a station in Greece. In the lower part of the figure, the responses of two ESDOF systems to the ground motion are shown, one corresponding to the undamaged and one severely damaged (ds_2) state. Note, that in the fragility analysis the same record is only used for two different damage states, if the record matches the response spectrum as defined by EC-8 in both systems' period ranges $[0.2T_1, 2.0T_1]$.

Determining the maximum drift values for each of the dynamic analyses yields a collection of PGA-drift pairs for a ESDOF system in a certain damage state (e.g., Figure C.4a). For each level of PGA, the ratio of number of exceedances with respect to the total number of simulations yields an estimate for the cumulative exceedance probability at this level of ground motion. Least squares fitting a lognormal CDF

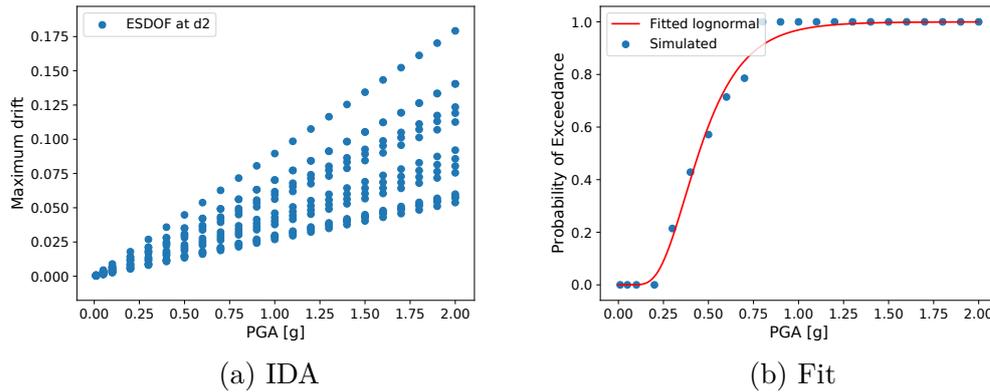


Figure C.4.: Example for a) maximum drift results for various ground motions (PGA) for a single ESDOF (308 simulations) and b) a fragility curve fit to these drift values for going from damage state ds_2 to the next damage state $ds_3 = 0.021$.

to these yields a probability model over the whole range of ground motions. Figure C.4 shows an example for a curve fit to simulation results.

C.4. Fragility for buildings in Nablus

Performing the EC-8 based capacity analysis, bilinear capacity curves with yielding and ultimate conditions as shown in Table 5.5 were estimated. The EC-2 design procedure used (Appendix C.1) seems to fail creating realistic larger building typologies (7 and 10 floors), as the derived capacities show unrealistically high values. Nevertheless, the models were not improved as this would require a more thorough analysis, beyond the scope of the demonstration purpose.

In order to obtain the dynamic building response for the buildings in Nablus, an IDA is performed with ground motion records from the Engineering Strong Motion (ESM) database (Luzi et al. 2015). Unfortunately, local recordings are few (18 records for 7 events), with no v_{s30} measurements for stations, with low magnitudes ($\leq M_w 5.5$), and recorded at large epicentral distances ($60\% \geq 80km$). Lacking sufficient local recordings, thus ground motions are selected from other regions available from the ESM database, matching the prescriptions of EC-8 for site-specific record selection.

Records with zero period spectral response larger than $a_g = 0.16g$ (PGA with 10% probability of exceedance in 50 years; Haas et al. (2016b) with Campbell and Bozorgnia (2014) GMPE) were selected. The v_{s30} values derived from topography (Wald and Allen 2007) suggest EC-8 soil types of type *B* (360-800 m/s) and few places of type *C* (180-360 m/s) in Nablus. Thus, two types of ground motion sets depending on the EC-8 soil type of the building are collected, for the period range of all ESDOF models (in any damage state) rounded to 0.05 s . Finally, the record selection

is constrained by magnitude and distance based on the disaggregation (Bazzurro and Cornell (1999); Figure C.5) of the hazard with 10% probability of exceedance within 50 years obtained from the same PSHA model using 0.5 M_w and 10 km distance bins and selecting events which have at least 0.1% chance to cause a ground motion exceedance of a_g ($M_w = 4.25 - 7.25$, $R_{epi} = 5 - 35km$). Note, that rather than deaggregating the rate (Bazzurro and Cornell 1999), the formulation used in OpenQuake deaggregates the exceedance probability. However, this yields similar results and is considered sufficient for the purpose here.

P(IM > PSHA(10% in 50y) | M,R)

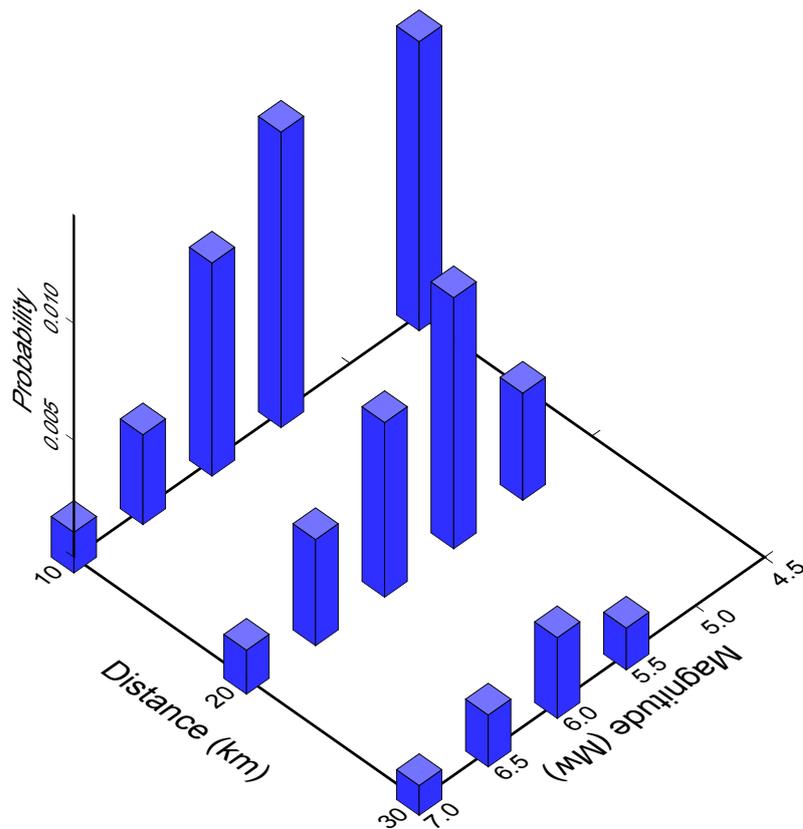


Figure C.5.: Disaggregation in 0.5 M_w and 10km distance (R) bins of the mean hazard curve on rock for Nablus obtained using the hazard model by Haas et al. (2016b) and a single GMPE (Campbell and Bozorgnia 2014). Only M-R combinations with at least 0.1 % chance of exceeding the mean hazard level with 475 years average return period are shown.

Using REXELite (Iervolino et al. 2011) with a lower and upper tolerance for the deviation from the response spectrum by 10% and 30%, respectively, the minimum required seven records (both horizontal components) are selected from the ESM database for each design period. This gives a ground motion set of 40 records for 30 events from Greece, Italy, and Turkey shown in Figure C.6.

Selecting for each ESDOF the ground motion set with the closest eigenperiod used for its composition, the IDA scales the ground motions to PGA levels of 0.01, 0.05,

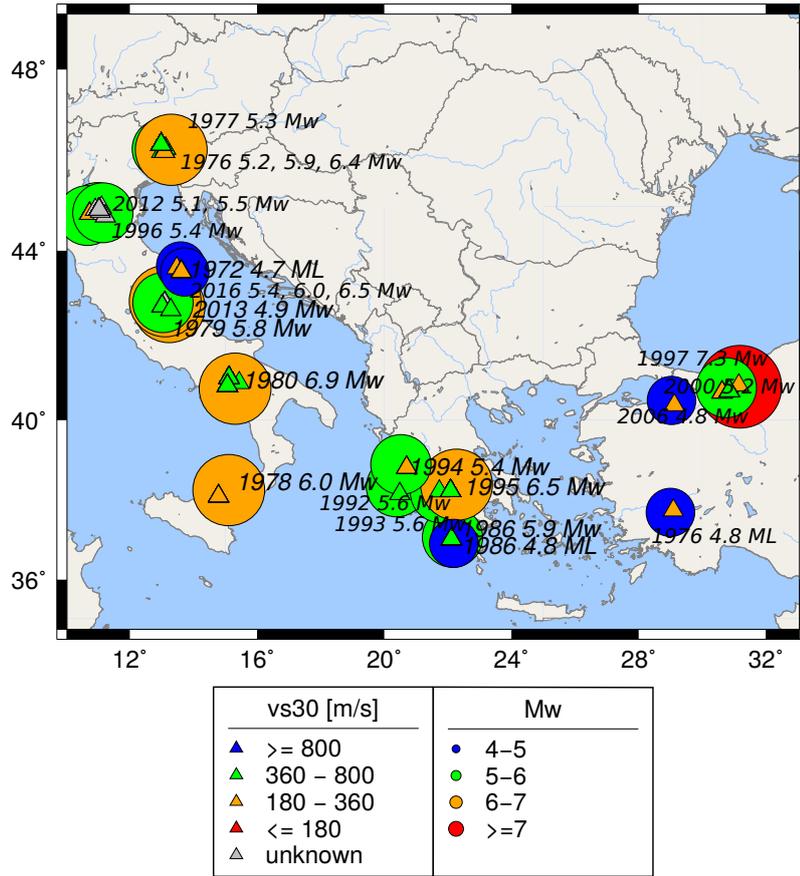


Figure C.6.: Recordings selected according to Eurocode-8 requirements from the ESM database (Luzi et al. 2015) employing REXELite Iervolino et al. (2011).

0.1, 0.2,..., 2.0 g , starting from the individual damage states, which results in 14 simulations for each level of ground motion and initial damage state.

Collecting the maximum drift levels for each level of PGA and fitting a lognormal distribution to these, gives the exceedance probability function of drift values for the respective next damage state d_i to d_{i+1} . Note, secant stiffness and increased damping for the ESDOF at corresponding drift values depend on the different starting damage states. During a single event a building might experience more than one damage state transition, i.e., deteriorate and change properties during the event. For reasons of simplicity, a much simplified approach is used by assuming independence of these transitions. This means all exceedance probabilities starting from a limit state ds_i whose transitions to ds_j ($j \neq i + 1$) could not be simulated with the linear model were calculated using equation 4.20. Table 5.7 lists the parameters for the estimated fragility functions.

C.5. Comparing fragility method to other approaches

In order to judge to which extent the determined fragility models are realistic several checks are performed. For the testing purpose, the applied methodology (see Appendices C.1, C.2, and C.3) was applied to the ICONS building (Pinto et al. 2002), a 4 storey reinforced concrete moment frame building, which was built and analyzed on a shaking table in the Institute for the Protection and the Security of the Citizen European Laboratory for Structural Assessment (ELSA) in Ispra, Italy during the ICONS project. The building has three frames, two with 5 meter and one with 2.5 meter spans, each story is 2.7 meters in height. Columns are rectangular and vary in size and reinforcement geometry. Figure C.7a shows the hysteretic behavior of the 4 stories of the structure as observed by Pinto et al. (2002) during the experiment. Next to it, Figure C.7b shows the capacity curves for both directions and the idealized bilinear model as obtained by applying the Eurocode-8 based approach outlined previously. They seem to capture the behavior sufficiently well.

Based on this capacity curve and the outlined approach an incremental dynamic analysis has been carried out, yielding the fragility models shown in Figures C.8a (soil type B) and C.8b (soil type C; both red curves). These curves model the transitions from the undamaged state to any other damage grade as defined by HAZUS. The gray lines in the background indicate corresponding fragility models for RC Moment Frame buildings (with different heights and design), available from the OpenQuake database (<https://platform.openquake.org/>; accessed May 2018) The proposed approach returns reasonable curves for the ICONS building.

Figure C.9 shows the increasing probabilities of exceedance for different starting damage states (only soil type B is shown). Please note that the stiffness and damping did not change going from undamaged to slight damage and thus the curves (except for the missing $d_{s,j} = 1$ for $d_{s,i} = 1$) are identical for the first damage state.

In the ICONS experiment (Pinto et al. 2002) the measured first modal frequency for the undamaged building was 1.5 Hz, the Rayleigh method applied here yields 0.9 Hz. The bare frame model used in the experiment almost collapsed at $0.29g$ (due to soft story failure in the third story) and the test was stopped and the model repaired and strengthened. The results of the subsequent tests can thus not be used for further comparison. The period elongation observed in the experiment till the near collapse was a factor of 3 whereas the secant stiffness used here yields a factor of 2.15 at d_u .

The fragility curves, can be compared to estimates which can be obtained using SPO2IDA (Vamvatsikos and Cornell 2006), an established tool that uses the typical shape of IDA curves to predict its results from a backbone capacity curve without actually running an IDA. Figure C.10 shows estimated fragilities for the (intact) ICONS building obtained with this program starting from the capacity curve determined using the EC-8 assessment. The resulting curves suggest that the ICONS building is less fragile as estimated with the approach used within this study.

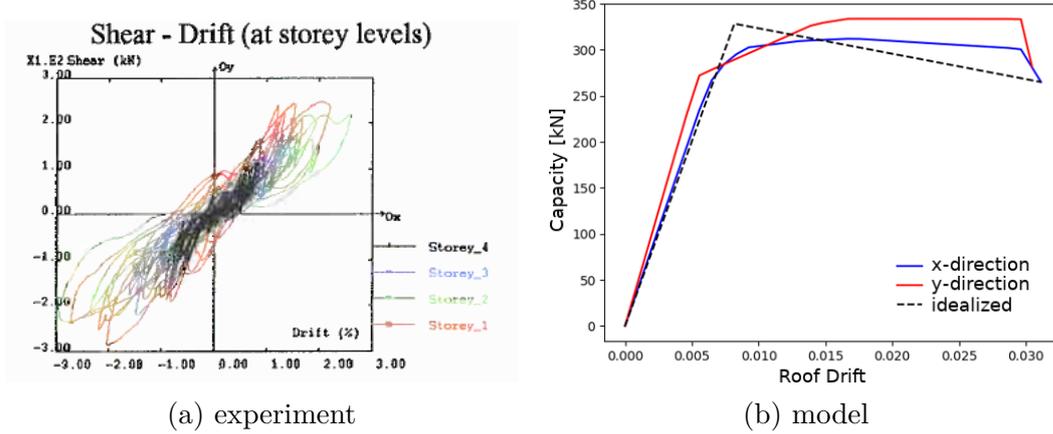


Figure C.7.: Hysteretic behavior observed during the experiment for an event with spectral acceleration of about 1.1 g at 0.4s by Pinto et al. (2002) (left), capacity curves resulting from the model developed here, in both directions, and the idealized bilinear capacity model (dashed line) fitted via least square residuals to the weaker directions' curve (right).

Another possible mean for comparison is to use a static analysis. In order to do so the capacity curve is first converted to spectral acceleration and displacement using the participation factor determined from the deformation shape ψ of the Rayleigh analysis and storey masses m_i :

$$\alpha = \frac{\sum_i \psi_i m_i}{\sum_i \psi_i^2 m_i} \quad (\text{C.40})$$

then the spectral acceleration is determined from:

$$S_{ae} = \frac{V_{base}}{\alpha \sum_i m_i g} \quad (\text{C.41})$$

and the spectral displacement as:

$$S_{de} = \frac{d_{roof}}{\alpha} \quad (\text{C.42})$$

For each of the defined limit states an intersection, i.e., performance point, between this spectral capacity curve and an elastic EC-8 response spectrum is determined. Non-linearity is accounted for using a ductility based reduction factor (Vidic et al. 1994):

$$R_\mu = \begin{cases} (\mu - 1) \frac{T}{T_c} & ; \text{ for } T < T_c \\ \mu & ; \text{ for } T \geq T_c \end{cases} \quad (\text{C.43})$$

on the response spectrum (or capacity) as:

$$S_a(T) = \frac{S_e(T)}{R_\mu} \quad (\text{C.44})$$

and:

$$S_d(T) = \frac{\mu}{R_\mu} S_d(T) \quad (\text{C.45})$$

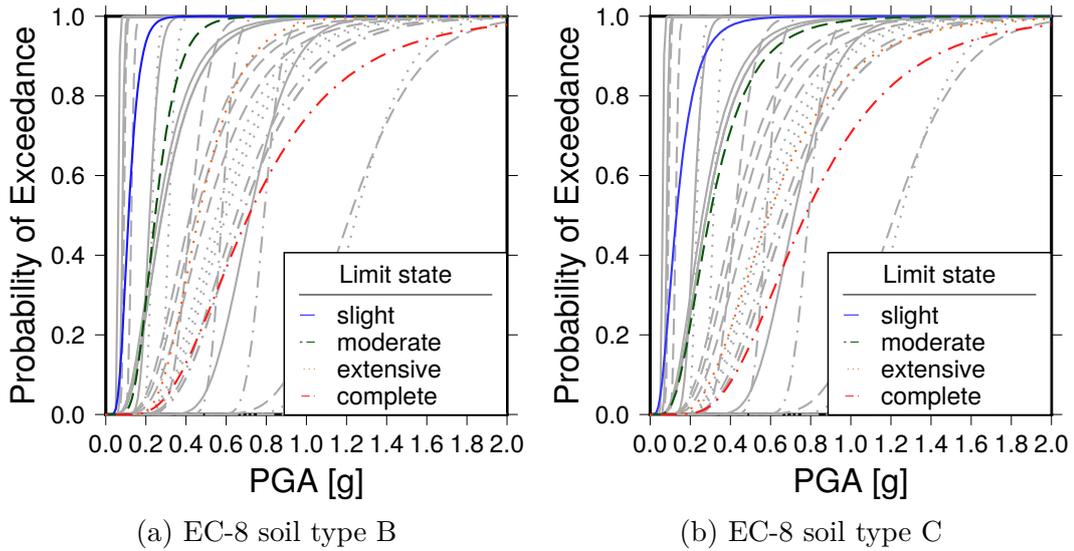


Figure C.8.: Fragility curves for the intact ICONS building on EC-8 soil types B and C. Gray curves indicate estimates for RC Moment Frame structures available from the OpenQuake database (<https://platform.openquake.org/>; accessed May 2018).

Assuming that the EC-8 response spectrum represents a median estimate for the ground motion spectrum, the level of zero period response (PGA) that yields the performance point EC-8 spectrum can be interpreted as the median value of the fragility curve. See Figure C.11 for such an estimate for the ICONS building. The variability could be determined from variations of empirical data such as found in the EMS-98, but is skipped here. Comparing the three resulting median values of the fragility curves for all three approaches, one observes that the simple dynamic approach presented in the previous sections and applied to the ICONS building here yield estimates in between the static and the SPO2IDA estimate.

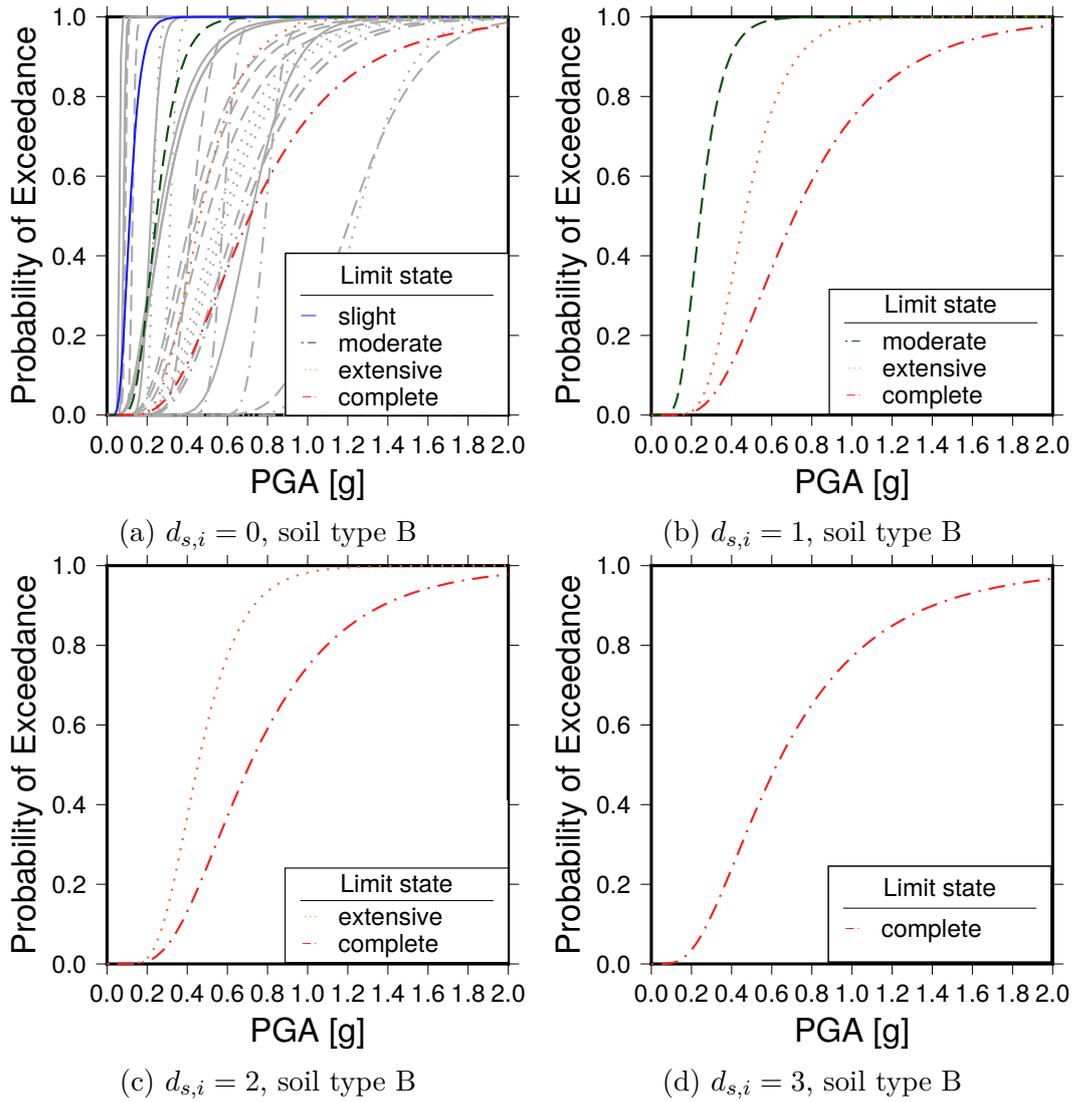


Figure C.9.: Increasing probability of exceedance for the ICONS buildings starting from different damage states on EC-8 soil type B.

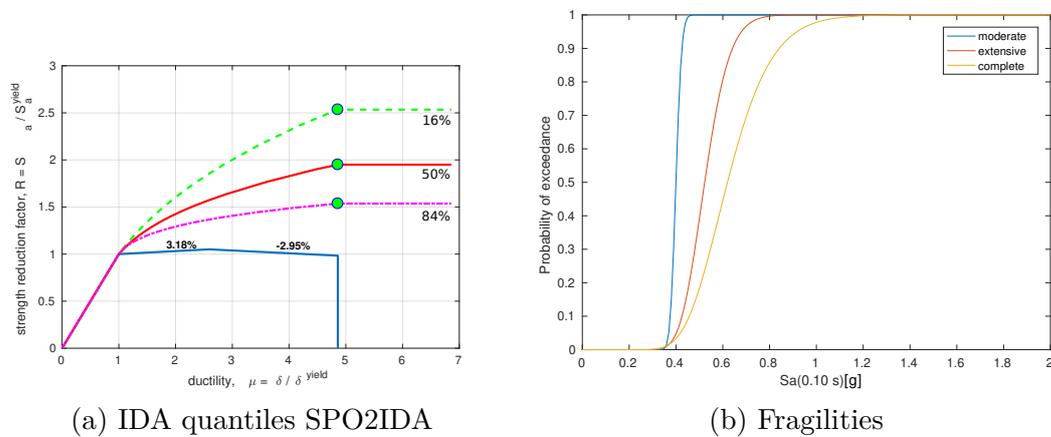


Figure C.10.: Fragility curves for the damage states *moderate*, *extensive*, and *collapse* for the ICONS building using SPO2IDA (Vamvatsikos and Cornell 2006)

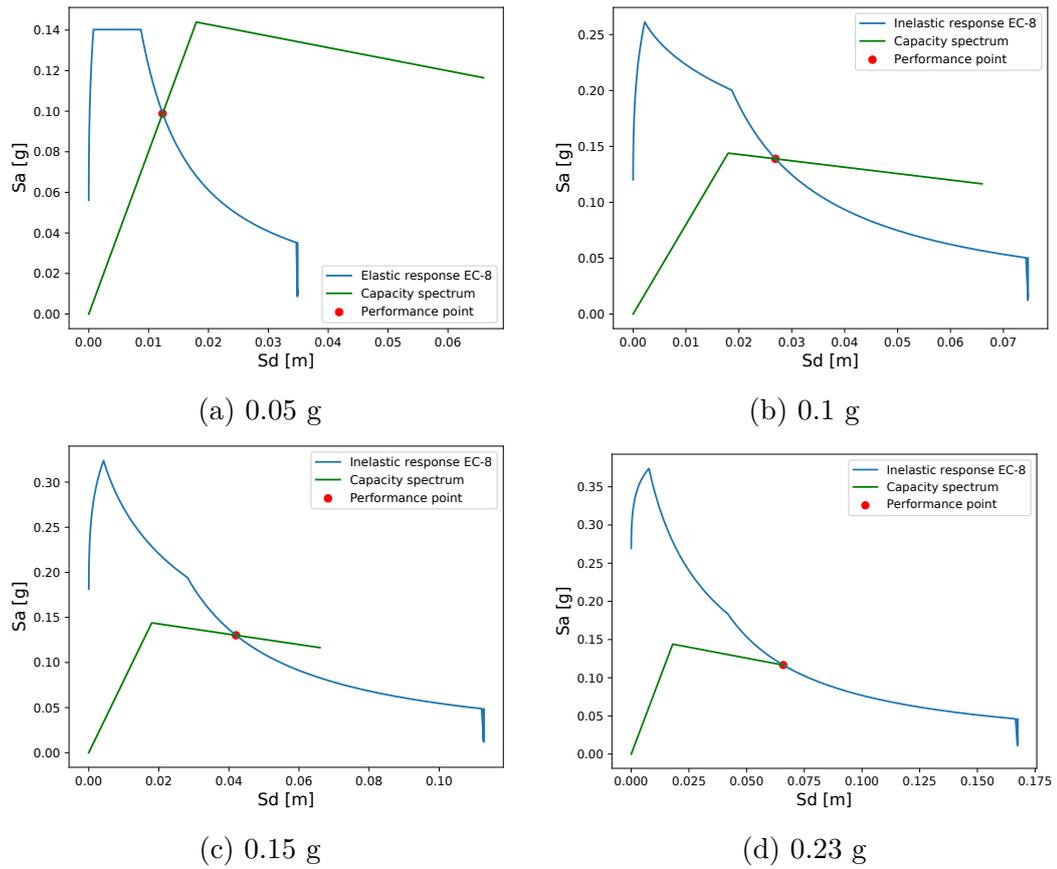


Figure C.11.: Performance points, determined at the spectral displacement values corresponding to the HAZUS damage states, using EC-8 elastic and inelastic response spectra, the latter depending on the ductility of the building capacity at the performance point.

D. Simulation input examples

Below are example input files for the simulation. Please note that they are shortened for readability and thus don't represent the full model used here.

D.1. Configuration file

```
[environment]
oq_path = /home/mhaas/openquake
oqdata = /home/mhaas/oqdata/db.sqlite3
parallel = True

[simulation]
#number of simulations
N=10000
#seed for simulation
seed = 42
#negative seed for etas code
etas_seed = -100
#fixed probability building repair 1:yes 0:no
# if yes: --> mu defines probability
# (e.g., mu=0: never repaired, mu=1: immediate repair)
fixed_repair=1
#state dependent fragilities (higher probability) 1:yes 0:no
cumulative_damage=1
#use previous hazard calculation 1:yes 0:no
previous_hazard=1
#session_id
previous_hazard_session=21

[input]
assets_file = Input/rrvs_exposure_model_0_999.csv
gmpe = CampbellBozorgnia2014
#factor to multiply values returned by gmpe to fragility values [g]
gmpe_unit = 1
intensity = "PGA":[0.01,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,
                  0.9,1.0,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2.0]
#depth and strike
depth_grid = Input/depth.csv
```

```

cmt_file = Input/cmt.csv

#file with fragility functions
fragility_file = Input/median_fragilities.csv
#parameters for lognormal repair function
#(in case of fixed_repair=1 mu defines fixed repair probability)
mu=1
sd=0.5

declustered_catalog = Input/Gruenthal_declustered_catalogue.csv
catalog = Input/DESERVE_Catalogue.csv

#maximum distance in degrees an event may have from assets
max_dist = 3

[smoothing]
#minimum magnitude
mmin=4.0
#catalog extent
extent = [24.55,29.95,37.80,40.80]
#discretization
dx,dy = 0.1,0.1
#years of completeness starting from mmin
completeness = [1987,1963,1945,1895,1870,1000,330]
#magnitude bins
mag_completeness = .5
#catalog end year
end = 2014
#magnitude discretization for frankel smoothing
mag_discr = 0.2
#b-value
b_val = 1.12
#smoothing width in km, truncated at 3 times this value
smoothing = 20.

[etas]
#catalog start year used as input catalog start
year_t=1915
#observation period for each simulation
T = 1.0
#maximum magnitude for GR
mmax = 8.0
#time delay of sequence onset in days
c = 0.003
#power law factor for rate decay with time
p = 1.1
#aftershocks added to simulation catalog yes:1 or no:0
aftershocks = 1

```

```
#real mainshocks, i.e., sequences are modeled
#but only mainshocks kept yes:1 or no:0
#NOTE: this only works if aftershocks = 1
#otherwise only Poissonian events are generated in the ETAS code
real_mainshocks = 0
```

D.2. Seismic catalog

```
year,month,day,hour,minute,second,longitude,latitude,depth,...
Mw,Source,sourceMagnitudeType
:
```

```
2014,10,11,19,14,32.81,35.8357,36.101,16.7,3.4,KOERI,ML
2014,10,12,6,37,1.01,30.5862,37.0678,3.5,3.4,KOERI,ML
2014,10,12,19,15,14.89,31.3773,36.0555,15.6,3.4,KOERI,ML
2014,10,14,9,10,54.86,30.2797,36.9812,4.5,3.6,KOERI,ML
2014,10,14,12,9,7.39,36.559,35.5963,12.6,3.4,KOERI,ML
2014,10,15,4,42,53.11,30.6052,37.0702,8.5,3.3,KOERI,ML
2014,10,15,11,38,15.73,30.581,37.0835,4.3,3.5,KOERI,ML
2014,10,16,6,10,35.56,30.9895,35.9132,6.9,4.2,NCEDCB,MB
2014,10,16,8,20,49.45,30.5888,37.0932,12.6,3.7,KOERI,ML
2014,10,16,12,40,10.08,30.9247,35.8354,57.94,4.3,NCEDCB,MB
2014,10,16,20,4,44.2,30.5732,37.0962,5,3.3,KOERI,ML
2014,10,16,21,42,40.27,34.5764,27.0825,1,3.4,GII,MD
2014,10,18,13,41,32.88,30.999,35.8296,10,4.2,NCEDCB,MB
2014,10,19,3,18,58.13,30.5718,37.1012,5,3.3,KOERI,ML
2014,10,19,17,47,59.796,34.5051,28.7622,1,3.4,GII,MD
2014,10,19,22,4,11.944,33.8917,34.8161,31,3.7,GII,MD
2014,10,20,21,54,0.62,30.9485,35.8358,9.1,4.2,NCEDCB,MB
2014,10,23,5,28,31.71,30.5215,37.1685,5.4,3.4,KOERI,ML
2014,10,25,10,43,33.89,37.1875,34.071,3,3.4,GII,MD
2014,11,1,13,33,52.2,33.923,35.0688,13.1,3.6,KOERI,ML
2014,11,7,17,20,5.28,34.6327,36.9227,5,3.7,KOERI,ML
2014,11,9,8,4,26.94,33.5855,34.641,4,3.5,KOERI,ML
2014,11,11,4,30,17.416,32.1904,33.7761,25,4.2,GII,MB
2014,11,12,16,25,34.4,31.48,35.93,14,4.3,NCEDCB,MB
2014,11,15,9,4,52.76,30.8935,36.407,2.4,3.7,KOERI,ML
2014,11,19,19,14,32.21,31.151,36.1032,12.4,3.3,KOERI,ML
2014,11,21,20,20,39.57,32.9937,33.7677,83.6,3.9,KOERI,ML
2014,11,23,7,31,6.77,30.9725,35.9488,11,3.3,KOERI,ML
2014,11,25,21,52,55.81,31.5562,35.2815,14.9,3.9,KOERI,ML
2014,12,7,8,49,1.96,34.5002,36.7643,5,3.9,KOERI,ML
2014,12,9,11,41,3.8,38.8483,35.9337,4.9,3.3,KOERI,ML
2014,12,12,20,21,14,31.1268,36.0157,10.3,3.3,KOERI,ML
```

2014,12,17,14,24,56.12,37.3397,37.3027,4.6,3.3,KOERI,ML
2014,12,24,17,59,49.09,30.3635,35.724,9.8,3.3,KOERI,ML
2014,12,24,18,11,35.72,30.3332,35.7435,7.8,4.3,NCEDCB,MB
2014,12,25,6,52,38.05,35.7743,37.6945,4.6,3.4,KOERI,ML
2014,12,25,14,7,50.21,34.2228,34.7749,9,4.1,GII,MD
2014,12,28,16,29,0.46,36.586,34.8192,5,3.7,KOERI,ML

D.3. Declustered catalog

day,depth,dtime,hour,latitude,longitude,magnitude,magnitudeType,...
minute,month,second,sigmaMagnitude,year

:

11,25,2014.860788,4,33.7761,32.1904,4.2,MB,30,11,17.416,0.26,2014
24,7.8,2014.980159,18,35.7435,30.3332,4.3,MB,11,12,35.72,0.26,2014
25,9,2014.982435,14,34.7749,34.2228,4.1,MD,7,12,50.21,0.3,2014

D.4. Depth model

lon,lat,2.5km,10km,22.5km,40km,
31.95,26.55,0.18,0.54,0.19,0.09
31.95,30.55,0.10,0.52,0.27,0.11
31.95,34.55,0.04,0.34,0.35,0.27
31.95,38.55,0.11,0.39,0.18,0.32

:

D.5. Source mechanisms

t,lat,lon,z,m,s,d,r

:

2010/11/14 23:08:25.0,36.58,36.01,2.5,5.0,24.0,53.0,-94.0
2012/01/30 17:04:28.8,27.79,33.99,28.4,4.6,343.0,33.0,-75.0
2012/05/11 18:48:30.2,34.3,34.12,20.5,5.4,112.0,66.0,-16.0
2012/07/22 09:26:03.0,37.55,36.38,7.6,5.0,38.0,53.0,-78.0
2012/09/19 09:17:46.8,37.31,37.1,7.0,4.9,210.0,48.0,-11.0
2012/10/19 03:35:10.9,32.54,30.98,9.5,5.1,336.0,58.0,57.0
2013/12/28 15:21:04.0,36.04,31.31,39.3,0.0,323.0,29.0,96.0

D.6. Assets

```
id,lon,lat,vs30,bt,p,dS,dt,occ,val,n
0,35.265,32.222,443.88,2f_1980_B,1,0,0,10,209089.23,1
1,35.231,32.237,502.11,4f_1980_B,1,0,0,10,599608.83,1
2,35.220,32.225,730.83,7f_2000_B,1,0,0,10,432839.87,1
3,35.273,32.213,760.00,2f_1980_B,1,0,0,10,484788.80,1
4,35.231,32.238,502.11,4f_1990-2000_B,1,0,0,10,859468.84,1
:
```

D.7. Fragilities

```
model_id,soil_type,di,dj,q,mu,sd,building_type
2f_1980,B,0,1,0.33,-2.06,0.68,0000000041_RCmodel_02f_1980_solid
2f_1980,B,0,1,0.66,-1.25,0.45,0000000025_RCmodel_02f_1980_ribbed2way
2f_1980,B,0,1,1.0,-1.24,0.47,0000000009_RCmodel_02f_1980_ribbed1way
2f_1980,B,0,2,0.33,-1.19,0.62,0000000041_RCmodel_02f_1980_solid
2f_1980,B,0,3,0.33,-0.73,0.59,0000000041_RCmodel_02f_1980_solid
2f_1980,B,0,2,0.66,-0.46,0.43,0000000025_RCmodel_02f_1980_ribbed2way
2f_1980,B,0,2,1.0,-0.44,0.45,0000000009_RCmodel_02f_1980_ribbed1way
2f_1980,B,0,4,0.33,-0.37,0.60,0000000041_RCmodel_02f_1980_solid
2f_1980,B,0,3,0.66,-0.18,0.44,0000000025_RCmodel_02f_1980_ribbed2way
2f_1980,B,0,3,1.0,-0.17,0.48,0000000009_RCmodel_02f_1980_ribbed1way
2f_1980,B,0,4,0.66,0.11,0.47,0000000025_RCmodel_02f_1980_ribbed2way
2f_1980,B,0,4,1.0,0.12,0.49,0000000009_RCmodel_02f_1980_ribbed1way
2f_1980,B,1,2,0.33,-1.29,0.68,0000000041_RCmodel_02f_1980_solid
2f_1980,B,1,3,0.33,-0.75,0.60,0000000041_RCmodel_02f_1980_solid
:
```

E. Computational costs

Figure E.1 shows the computation time, depending on the number of 50 year simulations, the number of computing threads, and the number of simulated buildings. This of course very limited benchmarking was performed on a Desktop computer with an Intel i5-4570 CPU with 4 cores with 3.20GHz, 4GB DDR3 RAM (1.6GHz) and a standard SATA hard drive. The computation time shows a linear increase with the number of simulations (left), but a non-linear increase with the number of buildings (right). Only a small computational overhead is observed in case of parallelization. In general the code was not thoroughly optimized for efficiency, though. Post-processing, i.e., converting damage grade transitions to losses and generating plots is handled separately and is not included in the computation time in Figure E.1. Depending on the number of simulations and buildings post-processing may take several minutes to several hours. As all routines (except plotting) are strictly parallelized, more powerful architectures likely yield a strong performance increase.

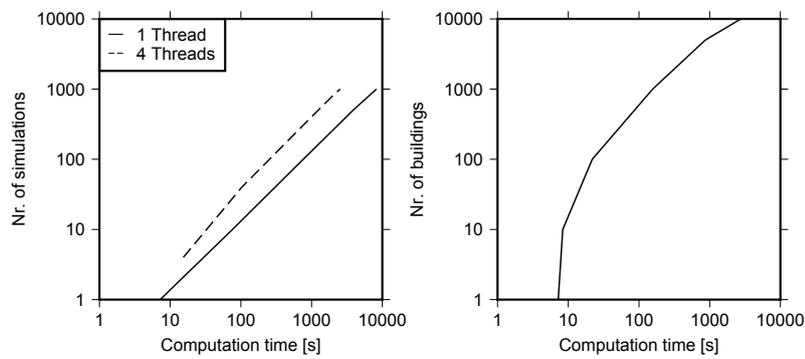


Figure E.1.: Computation time depending on the number of simulations and number of threads (left) and number of buildings (right). While an increase of simulations increases the computation time linearly, increasing building number causes a non-linear increase. The number of threads linearly reduces the computational time with a small overhead, which is only relevant for small numbers of simulations.

F. Statistical tests

In addition to the Kolmogorov-Smirnov test described in the section 6.3, additional one-sided tests can be performed on model pairs. Another non-parametric test is the Mann-Whitney-U-test (also Wilcoxon sum of ranks; Wilcoxon 1945; Mann and Whitney 1947). The test assumes that both samples have the same distribution, but are potentially shifted by a factor a in the underlying distribution $f_1(x) = f_2(x) + a$. As only one sided testing will be considered here, thus the null hypothesis for the test is $H0 : a = 0$ and the alternative hypothesis is $H1 : a > 0$. For the test, the observations for both models are pooled and ranked, where for identical ranks the median rank is assigned. For each of the models a sum of ranks (R_1, R_2) and a test statistics as:

$$U_i = n_1 n_2 + n_i(n_i + 1)/2 - R_i \quad (\text{F.1})$$

are calculated.

The null hypothesis is rejected if the minimal value $\min(U_1, U_2)$ is below a critical value $U_{critical}$, which is approximately normally distributed. Estimating a value z as:

$$z = \frac{U - n_1 n_2 / 2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12}} \approx N(0; 1) \quad (\text{F.2})$$

and given a significance level α the null hypothesis is rejected if $1 - F(z | \mu = 0; \sigma = 1) < \alpha$ where F denotes the CDF of the standard normal distribution.

The Mann-Whitney-U-test requires that the variances are identical. Thus, a test as the Brown-Forsythe test (Brown and Forsythe 1974) can be used to determine if the variance for (non-normal) distributions is different. For two samples, the test statistic for the latter is calculated as:

$$BF = \frac{\sum_{i=1}^2 n_i (\tilde{Y}_i - \tilde{Y})^2}{\frac{1}{\sum_{i=1}^2 n_i - j} \sum_{i=1}^2 \sum_{j=1}^n (Y_{i,j} - \tilde{Y}_i)^2} \quad (\text{F.3})$$

with Y the median of all samples, Y_i the sample median per model and Y_{ij} the individual samples deviation of the sample median:

$$Y_{ij} = |X_{ij} - \tilde{X}_i| \quad (\text{F.4})$$

with \tilde{X}_i the median of the samples for model i .

The test statistic is approximately F-distributed with $df_1 = m - 1$ and $df_2 = (n_1 + n_2) - k$ degrees of freedom. The null hypothesis is rejected if $1 - F(BF | df_1 = 1, df_2 = 199) < \alpha$, where F is the CDF of the F-distribution.

If the Brown-Forsythe test fails, the distributions have variance heterogeneity and the Mann-Whitney U-test can not be applied to test the significant of observed differences. A test for variance heterogeneous distributions is the Welch's test (Welch 1947), a variant of the Student t-test. This test however requires normality of the distributions, which is not necessarily the case here and the distributions have thus to be transformed to a scale were they have normal shape. A common transformation for this is the Box-Cox (Box and Cox 1964) transformation, which transforms a random variable as:

$$Y^{(\lambda)} = \begin{cases} \ln(Y + c) & ; \text{if } \lambda = 0 \\ \frac{(Y+c)^\lambda - 1}{\lambda} & ; \text{else} \end{cases} \quad (\text{F.5})$$

After the transformation the distributions are approximately normal and the Welch's t-test can be applied. The test statistic for the two distributions is calculated as:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \quad (\text{F.6})$$

with \bar{X}_1 , s_1^2 , \bar{X}_2 , and s_2^2 mean and variance of the two sample distributions. In addition, the degrees of freedom resulting from the variance have to be determined as:

$$\nu \approx \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{s_1^4/(n_1^2\nu_1) + s_2^4/(n_2^2\nu_2)} \quad (\text{F.7})$$

where $\nu_i = n_i - 1$ is the degrees of freedom of the variance for distribution $i = 1, 2$.

The test statistic is approximately t-distributed with ν , the null hypothesis is rejected if $1 - F(t|\nu) < \alpha$ where F is the CDF of the t-distribution.

Table F.1 shows all tests and all model combinations for 50 years and Table F.2 those for 1 year simulations. In most cases, the tests agree on the results of the Kolmogorov-Smirnov test. They only disagree on significance, for the following: in the boundary case, also previously discussed in Figure 6.9, the loss distributions at 5% and 10% PoE for the *ETAS* and *ETAS+DR(1,0.5)+SD* models, at 5 % PoE for the *DR(1,0.5)* and *FR(1)* models, and at 5% PoE the *ETAS+DR(-0.1,0.4)+SD* and 50.

Table F.1.: Statistical test for pairwise model combinations with higher and lower median values at different levels of exceedance probability (PoE[%]). Lavene (Brown-Forsythe; statistic: L and p-value: p_L) test for variance inhomogeneity, Mann-Whitney-U-test for significant shift in distributions (U , p_U), in case of variance heterogeneity t-test on Box Cox transformed samples (T , p_T), Kolmogorov-Smirnov test (KS , p_{KS}) and size of effect (d) for the mean values.

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
50y ETAS+DR(1,0.5)+SD	50y FR(1)	1	265	0.00			76.46	0.00	0.92	0.00	3.09
50y ETAS+DR(1,0.5)+SD	50y FR(1)	2	156	0.00			84.03	0.00	0.96	0.00	3.63
50y ETAS+DR(1,0.5)+SD	50y FR(1)	5	100	0.00			95.10	0.00	0.97	0.00	4.12
50y ETAS+DR(1,0.5)+SD	50y FR(1)	10	119	0.00			136.11	0.00	1.00	0.00	5.88
50y ETAS+DR(-0.1,0.4)+SD	50y FR(1)	1	283	0.00			75.55	0.00	0.93	0.00	3.01
50y ETAS+DR(-0.1,0.4)+SD	50y FR(1)	2	180	0.00			81.18	0.00	0.94	0.00	3.46
50y ETAS+DR(-0.1,0.4)+SD	50y FR(1)	5	125	0.00			93.38	0.00	0.97	0.00	4.03
50y ETAS+DR(-0.1,0.4)+SD	50y FR(1)	10	120	0.00			133.46	0.00	1.00	0.00	5.74
50y ETAS+DR(1.3,0.5)+SD	50y ETAS+DR(1,0.5)	1	18	0.00			17.55	0.00	0.34	0.00	0.61
50y ETAS+DR(1.3,0.5)+SD	50y ETAS+DR(1,0.5)	2	10	0.00			14.10	0.00	0.29	0.00	0.63
50y ETAS+DR(1.3,0.5)+SD	50y ETAS+DR(1,0.5)	5	20	0.00			13.92	0.00	0.26	0.00	0.62
50y ETAS+DR(1.3,0.5)+SD	50y ETAS+DR(1,0.5)	10	2	0.16	365562	0.00			0.21	0.00	0.48
50y DR(1,0.5)+SD	50y FR(1)	1	0	0.93	503952	0.62			0.04	0.53	0.01
50y DR(1,0.5)+SD	50y FR(1)	2	0	0.92	502554	0.58			0.02	0.99	0.01
50y FR(1)	50y DR(1,0.5)+SD	5	0	0.97	495450	0.36			0.03	0.79	0.01
50y FR(1)	50y DR(1,0.5)+SD	10	0	0.96	494972	0.35			0.03	0.75	0.01
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)	1	125	0.00			26.89	0.00	0.50	0.00	1.02
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)	2	17	0.00			21.11	0.00	0.42	0.00	0.95
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)	5	28	0.00			18.94	0.00	0.35	0.00	0.83
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)	10	9	0.00			16.28	0.00	0.31	0.00	0.72
50y ETAS+DR(1,0.5)	50y DR(1,0.5)	1	46	0.00			69.74	0.00	0.92	0.00	2.92
50y ETAS+DR(1,0.5)	50y DR(1,0.5)	2	87	0.00			84.31	0.00	0.95	0.00	3.66
50y ETAS+DR(1,0.5)	50y DR(1,0.5)	5	59	0.00			82.12	0.00	0.95	0.00	3.59
50y ETAS+DR(1,0.5)	50y DR(1,0.5)	10	123	0.00			111.53	0.00	0.99	0.00	4.81

Continued on the next page

Table F.1 – continued from previous page

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(-0.1,0.4)+SD	1	1	0.34	500418	0.51			0.05	0.21	0.00
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)+SD	2	2	0.14	506032	0.68			0.04	0.46	0.01
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)+SD	5	1	0.36	508536	0.75			0.04	0.28	0.03
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(-0.1,0.4)+SD	10	0	0.91	470378	0.01			0.06	0.05	0.09
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)+SD	1	122	0.00			76.60	0.00	0.94	0.00	3.10
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)+SD	2	124	0.00			90.07	0.00	0.96	0.00	3.93
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)+SD	5	112	0.00			89.72	0.00	0.96	0.00	3.86
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)+SD	10	165	0.00			126.86	0.00	0.99	0.00	5.39
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)+SD	1	282	0.00			75.62	0.00	0.93	0.00	3.01
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)+SD	2	183	0.00			81.34	0.00	0.94	0.00	3.46
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)+SD	5	124	0.00			93.61	0.00	0.97	0.00	4.04
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)+SD	10	120	0.00			133.43	0.00	1.00	0.00	5.75
50y ETAS+DR(1.3,0.5)+SD	50y FR(1)	1	123	0.00			76.53	0.00	0.94	0.00	3.10
50y ETAS+DR(1.3,0.5)+SD	50y FR(1)	2	122	0.00			89.90	0.00	0.96	0.00	3.92
50y ETAS+DR(1.3,0.5)+SD	50y FR(1)	5	112	0.00			89.48	0.00	0.96	0.00	3.85
50y ETAS+DR(1.3,0.5)+SD	50y FR(1)	10	165	0.00			126.89	0.00	0.99	0.00	5.39
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)	1	84	0.00			16.18	0.00	0.34	0.00	0.71
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)	2	31	0.00			8.39	0.00	0.21	0.00	0.39
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)	5	24	0.00			18.25	0.00	0.36	0.00	0.82
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1,0.5)	10	0	0.67	324309	0.00			0.32	0.00	0.62
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1.3,0.5)+SD	1	14	0.00			2.61	0.00	0.16	0.00	0.13
50y ETAS+DR(1.3,0.5)+SD	50y ETAS+DR(1,0.5)+SD	2	1	0.26	438296	0.00			0.15	0.00	0.20
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1.3,0.5)+SD	5	0	0.66	427579	0.00			0.18	0.00	0.22
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1.3,0.5)+SD	10	4	0.05	426890	0.00			0.16	0.00	0.22
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)+SD	1	24	0.00			8.83	0.00	0.16	0.00	0.41
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)+SD	2	0	0.86	347496	0.00			0.24	0.00	0.52
50y ETAS+DR(1,0.5)+SD	50y ETAS+FR(1)	5	2	0.22	487285	0.16			0.09	0.00	0.02
50y ETAS+FR(1)	50y ETAS+DR(1,0.5)+SD	10	12	0.00			0.73	0.23	0.08	0.00	0.04

Continued on the next page

Table F.1 – continued from previous page

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)	1	268	0.00			80.10	0.00	0.94	0.00	3.24
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)	2	202	0.00			91.27	0.00	0.96	0.00	3.90
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)	5	128	0.00			97.37	0.00	0.97	0.00	4.20
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)	10	110	0.00			130.61	0.00	1.00	0.00	5.69
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1,0.5)	1	71	0.00			16.64	0.00	0.35	0.00	0.72
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1,0.5)	2	18	0.00			9.16	0.00	0.22	0.00	0.41
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1,0.5)	5	16	0.00			19.34	0.00	0.38	0.00	0.86
50y ETAS+DR(1,0.5)+SD	50y ETAS+DR(1,0.5)	10	0	0.59	299453	0.00			0.35	0.00	0.72
50y ETAS+DR(1,0.5)	50y FR(1)	1	46	0.00			65.67	0.00	0.90	0.00	2.74
50y ETAS+DR(1,0.5)	50y FR(1)	2	58	0.00			77.04	0.00	0.92	0.00	3.38
50y ETAS+DR(1,0.5)	50y FR(1)	5	39	0.00			79.98	0.00	0.95	0.00	3.50
50y ETAS+DR(1,0.5)	50y FR(1)	10	132	0.00			116.81	0.00	0.99	0.00	5.00
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1.3,0.5)+SD	1	21	0.00			2.49	0.01	0.13	0.00	0.13
50y ETAS+DR(1.3,0.5)+SD	50y ETAS+DR(-0.1,0.4)+SD	2	7	0.01			4.88	0.00	0.15	0.00	0.21
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1.3,0.5)+SD	5	0	0.63	439509	0.00			0.16	0.00	0.19
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+DR(1.3,0.5)+SD	10	3	0.07	454004	0.00			0.13	0.00	0.13
50y ETAS+FR(1)	50y DR(1,0.5)+SD	1	273	0.00			80.36	0.00	0.94	0.00	2.86
50y ETAS+FR(1)	50y DR(1,0.5)+SD	2	137	0.00			95.31	0.00	0.96	0.00	4.12
50y ETAS+FR(1)	50y DR(1,0.5)+SD	5	134	0.00			94.56	0.00	0.98	0.00	4.06
50y ETAS+FR(1)	50y DR(1,0.5)+SD	10	217	0.00			131.78	0.00	1.00	0.00	5.57
50y DR(1,0.5)+SD	50y DR(1,0.5)	1	0	0.93	448275	0.00			0.13	0.00	0.19
50y DR(1,0.5)+SD	50y DR(1,0.5)	2	3	0.07	426040	0.00			0.13	0.00	0.26
50y DR(1,0.5)	50y DR(1,0.5)+SD	5	3	0.11	500892	0.53			0.06	0.05	0.02
50y DR(1,0.5)	50y DR(1,0.5)+SD	10	0	0.69	435502	0.00			0.12	0.00	0.20
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)	1	287	0.00			79.13	0.00	0.94	0.00	3.15
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)	2	226	0.00			88.23	0.00	0.96	0.00	3.70
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)	5	156	0.00			95.60	0.00	0.97	0.00	4.11
50y ETAS+DR(-0.1,0.4)+SD	50y DR(1,0.5)	10	111	0.00			127.99	0.00	1.00	0.00	5.55

Continued on the next page

Table F.1 – continued from previous page

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
50y FR(1)	50y DR(1,0.5)	1	0	0.86	445289	0.00			0.14	0.00	0.20
50y FR(1)	50y DR(1,0.5)	2	4	0.06	424323	0.00			0.13	0.00	0.27
50y DR(1,0.5)	50y FR(1)	5	3	0.11	504282	0.63			0.07	0.01	0.03
50y DR(1,0.5)	50y FR(1)	10	0	0.73	438988	0.00			0.12	0.00	0.19
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)+SD	1	264	0.00			76.53	0.00	0.92	0.00	3.09
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)+SD	2	159	0.00			84.20	0.00	0.96	0.00	3.64
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)+SD	5	100	0.00			95.34	0.00	0.97	0.00	4.13
50y ETAS+DR(1,0.5)+SD	50y DR(1,0.5)+SD	10	119	0.00			136.07	0.00	1.00	0.00	5.89
50y ETAS+DR(1,0.5)	50y DR(1,0.5)+SD	1	45	0.00			65.75	0.00	0.90	0.00	2.74
50y ETAS+DR(1,0.5)	50y DR(1,0.5)+SD	2	60	0.00			77.22	0.00	0.93	0.00	3.39
50y ETAS+DR(1,0.5)	50y DR(1,0.5)+SD	5	39	0.00			80.23	0.00	0.94	0.00	3.51
50y ETAS+DR(1,0.5)	50y DR(1,0.5)+SD	10	132	0.00			116.81	0.00	0.99	0.00	5.01
50y ETAS+FR(1)	50y DR(1,0.5)	1	274	0.00			83.48	0.00	0.95	0.00	2.97
50y ETAS+FR(1)	50y DR(1,0.5)	2	172	0.00			102.45	0.00	0.97	0.00	4.38
50y ETAS+FR(1)	50y DR(1,0.5)	5	167	0.00			96.60	0.00	0.98	0.00	4.13
50y ETAS+FR(1)	50y DR(1,0.5)	10	205	0.00			126.46	0.00	1.00	0.00	5.39
50y ETAS+FR(1)	50y FR(1)	1	274	0.00			80.32	0.00	0.94	0.00	2.86
50y ETAS+FR(1)	50y FR(1)	2	134	0.00			95.14	0.00	0.96	0.00	4.12
50y ETAS+FR(1)	50y FR(1)	5	134	0.00			94.33	0.00	0.98	0.00	4.05
50y ETAS+FR(1)	50y FR(1)	10	217	0.00			131.81	0.00	1.00	0.00	5.57
50y ETAS+FR(1)	50y ETAS+DR(-0.1,0.4)+SD	1	17	0.00			8.82	0.00	0.20	0.00	0.40
50y ETAS+FR(1)	50y ETAS+DR(-0.1,0.4)+SD	2	2	0.22	345467	0.00			0.25	0.00	0.52
50y ETAS+DR(-0.1,0.4)+SD	50y ETAS+FR(1)	5	0	0.76	497032	0.41			0.07	0.02	0.02
50y ETAS+FR(1)	50y ETAS+DR(-0.1,0.4)+SD	10	11	0.00			2.68	0.00	0.09	0.00	0.13
50y ETAS+FR(1)	50y ETAS+DR(1.3,0.5)+SD	1	59	0.00			11.26	0.00	0.29	0.00	0.53
50y ETAS+FR(1)	50y ETAS+DR(1.3,0.5)+SD	2	2	0.22	400879	0.00			0.20	0.00	0.33
50y ETAS+FR(1)	50y ETAS+DR(1.3,0.5)+SD	5	1	0.42	441181	0.00			0.11	0.00	0.20
50y ETAS+FR(1)	50y ETAS+DR(1.3,0.5)+SD	10	2	0.13	430165	0.00			0.12	0.00	0.25

Continued on the next page

Table F.1 – continued from previous page

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)	1	124	0.00			80.29	0.00	0.95	0.00	3.25
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)	2	162	0.00			97.27	0.00	0.97	0.00	4.19
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)	5	142	0.00			91.66	0.00	0.97	0.00	3.93
50y ETAS+DR(1.3,0.5)+SD	50y DR(1,0.5)	10	155	0.00			121.55	0.00	0.99	0.00	5.20

Table F.2.: Statistical test for pairwise model combinations with higher and lower median values at different levels of exceedance probability (PoE[%]). Lavene (Brown-Forsythe; statistic: L and p-value: p_L) test for variance inhomogeneity, Mann-Whitney-U-test for significant shift in distributions (U , p_U), in case of variance heterogeneity t-test on Box Cox transformed samples (T , p_T), Kolmogorov-Smirnov test (KS , p_{KS}) and size of effect (d) for the mean values.

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
1y ETAS+FR(0.0000623)+SD	1y FR(1)	5	1	0.41	444738	0.00			0.09	0.00	0.19
1y ETAS+FR(0.0000623)+SD	1y FR(1)	10	1	0.32	442564	0.00			0.10	0.00	0.17
1y ETAS+FR(0.0000623)+SD	1y FR(1)	20	50	0.00			11.80	0.00	0.26	0.00	0.54
1y ETAS+FR(0.0)+SD	1y FR(1)	5	0	0.92	453803	0.00			0.09	0.00	0.15
1y ETAS+FR(0.0)+SD	1y FR(1)	10	1	0.25	489822	0.22			0.04	0.31	0.03
1y ETAS+FR(0.0)+SD	1y FR(1)	20	13	0.00			6.72	0.00	0.17	0.00	0.31
1y ETAS+FR(0.0000623)+SD	1y ETAS+FR(0.0)+SD	5	1	0.47	490455	0.23			0.03	0.68	0.05
1y ETAS+FR(0.0000623)+SD	1y ETAS+FR(0.0)+SD	10	0	0.82	451747	0.00			0.10	0.00	0.14
1y ETAS+FR(0.0000623)+SD	1y ETAS+FR(0.0)+SD	20	10	0.00			5.00	0.00	0.14	0.00	0.23
1y ETAS+DR(1,0.5)+SD	1y FR(1)	5	0	0.66	447200	0.00			0.10	0.00	0.17
1y ETAS+DR(1,0.5)+SD	1y FR(1)	10	0	0.95	451740	0.00			0.11	0.00	0.16
1y ETAS+DR(1,0.5)+SD	1y FR(1)	20	2	0.18	442825	0.00			0.13	0.00	0.20
1y ETAS+FR(0.0)+SD	1y real MS+ FR(1)	5	1	0.37	437340	0.00			0.11	0.00	0.22
1y ETAS+FR(0.0)+SD	1y real MS+ FR(1)	10	8	0.00			0.89	0.19	0.05	0.13	0.06
1y ETAS+FR(0.0)+SD	1y real MS+ FR(1)	20	6	0.02			2.04	0.02	0.10	0.00	0.10
1y ETAS+FR(0.0000623)+SD	1y real MS+ FR(1)	5	3	0.10	429157	0.00			0.11	0.00	0.27
1y ETAS+FR(0.0000623)+SD	1y real MS+ FR(1)	10	11	0.00			4.49	0.00	0.13	0.00	0.21
1y ETAS+FR(0.0000623)+SD	1y real MS+ FR(1)	20	32	0.00			7.14	0.00	0.20	0.00	0.33
1y ETAS+DR(1,0.5)+SD	1y real MS+ FR(1)	5	2	0.21	431455	0.00			0.10	0.00	0.25
1y ETAS+DR(1,0.5)+SD	1y real MS+ FR(1)	10	17	0.00			4.06	0.00	0.11	0.00	0.20
1y ETAS+DR(1,0.5)+SD	1y real MS+ FR(1)	20	0	0.98	501706	0.55			0.06	0.07	0.02
1y FR(1)	1y real MS+ FR(1)	5	1	0.43	484654	0.12			0.04	0.31	0.07
1y real MS+ FR(1)	1y FR(1)	10	17	0.00			-0.10	0.46	0.06	0.07	0.03
1y real MS+ FR(1)	1y FR(1)	20	2	0.19	439941	0.00			0.10	0.00	0.22

Continued on next page

Table F.2 – continued from previous page

Model high median	Model low median	PoE	L	p_L	U	p_U	T	p_T	KS	p_{KS}	d
1y ETAS+FR(0.0)+SD	1y ETAS+DR(1,0.5)+SD	5	0	0.74	505318	0.66			0.04	0.31	0.02
1y ETAS+DR(1,0.5)+SD	1y ETAS+FR(0.0)+SD	10	1	0.27	459867	0.00			0.10	0.00	0.14
1y ETAS+FR(0.0)+SD	1y ETAS+DR(1,0.5)+SD	20	5	0.02			2.50	0.01	0.08	0.00	0.12
1y ETAS+FR(0.0000623)+SD	1y ETAS+DR(1,0.5)+SD	5	0	0.68	496175	0.38			0.03	0.79	0.02
1y ETAS+DR(1,0.5)+SD	1y ETAS+FR(0.0000623)+SD	10	1	0.34	508993	0.76			0.05	0.09	0.00
1y ETAS+FR(0.0000623)+SD	1y ETAS+DR(1,0.5)+SD	20	32	0.00			7.58	0.00	0.20	0.00	0.35

G. Glossary

Table G.1.: Glossary of terms used that are not defined within the body of the main text.

Term	Description
adobe	Fired clay bricks.
agent-based models	Models that consider individual agents (e.g., persons) that interact with each other.
aggregate	Unite units by a criteria (e.g., spatial proximity).
aleatory uncertainty	Uncertainty resulting from the stochastic nature of a process.
alluvium	Sediments deposited by a river or lake.
ancillary data	Data that allows one to deduce information that is not directly contained.
APSHA	Probabilistic seismic hazard analysis, including ground motion from aftershocks. The result is stationary in time.
artifact	Observation that is artificially induced and not an actual feature.
Bayesian estimator	Estimates a parameter from observed data, considering a prior distribution based on additional knowledge about the parameter
bays of a building	A single frame in a moment frame building.
bootstrap	Random sampling with replacement from a distribution.
braced frame	Frame strengthened by diagonal elements.
brittle failure	Capacity loss due to compressional forces.
building irregularities	Structural irregularities that may strongly affect a buildings response, usually for the worse.
cantilever	Beam that is supported only on one side.
capacity curve	Curve that describes the envelope of the strength-deformation relation (hysteresis) of an element subject to (lateral) loading.
cascading events	Events which trigger additional events and may have been triggered themselves.
categorical variable	Variable that describes something qualitatively, hence cannot be ordered.

Continued on next page

Table G.1 – continued from previous page

Term	Description
census	Collection of data for a full population (no sampling).
centroid moment tensor	Description of the tension, compression, and shearing moments acting at the centroid of an earthquake during rupture in all three dimensions.
chord rotation	Angle between the perpendicular direction and the actual direction of a column, for small angles equal to drift.
classification	Assigning defined class labels to objects based on their attributes.
clay	Soil material composed mostly by clay minerals.
cleaned stone	Cut stone.
clustering	Unsupervised (data-driven) grouping of objects based on their attributes and using some metric of similarity among them.
cohesionless soil	Soil that is not cemented.
cohesive soil	Soil that is cemented.
composite	Material composed of several materials.
compression	Inwards oriented stresses.
compressional reinforcement	Reinforcement in the part of a column or beam which is under compression.
conservative choice	A choice that is more likely to overestimate the impact of something than to underestimate it.
convolution	Overlay of signals, mathematically a combination of multiplication and integration.
cumulative density function (CDF)	Integral over the probability density function, describes cumulative probabilities (equal or less).
correlation	Strength of the relationship between two variables.
crowd sourced	Data collected by a large number of volunteers (crowd).
crustal	Related to the Earth's crust.
damping	Effect that decreases vibration amplitudes over time.
declustering	Identifying earthquake clusters in a seismic catalog and the removal of fore- and after-shocks, so that only independent main shocks remain.
degree of freedom	Number of independent parameters of a process or system that can vary.

Continued on next page

Table G.1 – continued from previous page

Term	Description
demand surge	increase of material and labour cost in case of a large peak in demand.
detached building	Building that has no adjoining buildings on its sides.
deterministic	Input parameters uniquely define result, no stochastic component.
diaphragm	Connection between lateral and vertical load resisting systems of a structure.
dip	Angle between the surface and the rupture plane.
disaggregation	Splitting the hazard integrals into the most relevant contributors.
drift	Top displacement with respect to the height.
ductile	Plastic deformation before rupturing.
dwelling	Housing unit for a household.
torsion eccentricity	Where the center of stiffness and gravitational center are not identical.
eigenfrequency	Fundamental frequency of a oscillator.
eigenperiod	Fundamental period of a oscillator.
elastic rebound theory	Due to plate tectonics, stresses build up on a fault, these deform both sides of a fault until stress is large enough to fracture the bonding (earthquake), and build up again.
elastic response spectrum	Elastic response of harmonic oscillators with various fundamental periods to ground motion.
empirical	Knowledge derived from observations.
epicenter	Geographical coordinate (on the Earth's surface) of an earthquake's rupture center.
epidemic	Process that "infects" other elements rapidly.
epistemic uncertainty	Uncertainty resulting from a lack of knowledge about a process.
equilibrium	Balance of two or more forces.
ergodic assumption	A process has the same behavior in time as in space; for GMPE development: because the time-series of data at a station is not sufficiently long, recordings from similar stations are used..
ESDOF	Equivalent single degree of freedom system used to approximate a multi-degree of freedom system.
exceedance probability (EP)	Probability to exceed a certain value; determined from the survival function of a PDF, equivalent to 1-CDF.

Continued on next page

Table G.1 – continued from previous page

Term	Description
exhaustive	All features are captured (potentially generalized; e.g., trees and grass are assigned to a class vegetation).
faceted taxonomy	Instead of a strictly hierarchical classification.
fault	Part of the rock material that has been ruptured previously by earthquakes and forms a weak part of the rock, where future earthquakes might occur.
feature	Attribute of an object.
Finite Element Modelling	A complex body is discretized into elements for which the mathematical equations of their physics are simpler. Boundary conditions transmit the effects over the mesh of elements through the whole body.
fractal	Shows symmetries on many scales, extreme case is a self-similar pattern, i.e., a pattern that repeats itself on any scale; the scaling is a fraction (not an integer), e.g., doubling the length of each side results in an object with 1.3 (or any other fraction) times its original area.
moment frame	Only the frame of columns and beams of a building that resist moments induced by loading.
fuzzy scoring	Here: assigning a score of similarity between attributes that is not defined by a single number, but a range of numbers (fuzzy).
Gaussian	Following a normal distribution.
geocoded	Object that is geographically referenced.
geometrical spreading	decrease of a ground motion pulse as it travels away from its origin.
geotechnical	Investigation of soil parameters using measurements.
gravel	Small pieces of rock material that have been shaped/eroded through transport by water.
hardening	Plastical deformation which causes no strength loss only a decrease in stiffness.
hydrofracturing	Use of liquid injections under high pressure into rock material to create or expand fractures and increase porosity and permeability.
hypocenter	The central point of an earthquake rupture in spherical coordinates (longitude, latitude, depth).

Continued on next page

Table G.1 – continued from previous page

Term	Description
hysteresis	Defromation response of an object to external forces (here, loading and unloading).
incompleteness	Incomplete seismic catalogs are missing events that are expected to have occurred; assuming constant rates of seismicity over long periods. They are usually identified through rate changes for events of a particular size.
induced seismicity	Events induced through the antropogenic activity, i.e., events that would not have occurred naturally; These are different from triggered events, which would have released stress at some point.
informal construction	Construction that was not authorized by the government and usually does not follow an engineering design.
inhomogeneous	Object or group of objects with strongly varying properties.
instrumental seismicity	Seismicity that is measured and recorded using seismometers.
intensity prediction equation (IPE)	Like a GMPE, but prediciting macroseismic intensity instead of a physical parameter of ground motion.
Inter-event residual	Residuals in the ground motion observed at a station for similar events.
interface event (earthquake)	Earthquakes that occur at the interface between two plates (usually only referred to as such in subduction zones).
Intra-event residual	Residuals in the ground motion observed at similar stations for the same event.
isotropic	Spatially uniform in all directions.
Landsat	Satellite series launched, owned, and maintained by NASA and USGS since the early 70s. Data is available free of cost. The most recent satellite is the Landsat 8, operational since 2013.
landslides	Mass movement of soil that can be triggered by an earthquake.
landuse	Purpose and use of a piece of land.
lap splice	Length over which reinforcement bars overlap (and are connected).
lateral reinforcement	Reinforcement along the lateral axis of an element (stirrups in columns and beams).

Continued on next page

Table G.1 – continued from previous page

Term	Description
Latin Hyper Cube sampling	Sampling intervals of a distribution rather than randomly, e.g., 3 samples, one from 0-33%, one from 33-66%, and one from 66-100% of a CDF.
lemmatizing	Removing the pre- and suffixes of a word, reducing it to its stem.
likelihood	Probability of observing a set of data given a probability density function.
limestone	Rock formed from sedimented calcium carbonate.
liquefaction	Process whereby soil loses its shear resistance and behaves like a liquid during shaking (usually requires highly saturated soil with loose sediments).
logic tree branches	Logic trees encode epistemic uncertainty, a branch is one possible model.
logic tree weights	weight that defines the probability with which a logic tree branch is chosen, to calculate a single result/instance of a process model.
longitudinal	Along the elongated axis of a body.
tensional reinforcement	Reinforcement in the part of a column or beam which is in tension.
longitudinal reinforcement	Reinforcement along the longitudinal axis of a beam or column.
machine learning	Training an algorithm to recognize a pattern automatically.
magnitude	Size of an earthquake. This can be determined in several ways, for example, from the seismogram amplitude and epicentral distance (local /Richter magnitude m_L), from body waves (m_b), from surface waves (M_S), from the rupture area and seismic moment (M_w).
mainshock	The event with the largest magnitude within a seismic sequence.
mantle convection	The main heat transfer mechanism in the Earth's mantle, the result of advection and conduction.
Markov chain	A process whose next step depends only on the current state of the system, not its history.
MDOF	Multi-degree-of-freedom system, that can be used to describe a buildings response.
mobile mapping	Collection of geocoded data using a vehicle.

Continued on next page

Table G.1 – continued from previous page

Term	Description
modes of vibration	Standing waves that are used to describe and decompose the complex vibrations of an oscillating system.
Monte Carlo sampling	Repeated random sampling from a probability distribution.
microzonation	Geospatial analysis of soil properties affecting ground motion over a large scale.
moment of force	Force applied on a lever and acting over some distance on a point.
moment of inertia	Second moment of mass of a body.
multi-spectral	Covering a broad part of the spectrum of light, using several spectral bands.
mutually exclusive	One excludes the other, i.e., no overlap.
Nakamura H/V	Ratio between the frequency spectra of a multi component ground motion recording, the horizontal (with respect to the ground) divided by the vertical component. Only the fundamental frequency of the ground can be reliably obtained with this technique, especially if only ambient seismic noise is used.
normal faulting	Faulting resulting from tectonic rifting.
normalized	When a set of values are divided by a single value or scaled to a range using some norm, such as an euclidean norm.
nucleation	Point of origin of a process.
occupancy	Use of a building, sometimes also referring to the (time-variant) number of occupants.
ordinal	Data that can be ordered (e.g., by size).
overhangs	Part of a structure that is vertically supported by cantilevers.
panoramic image	An image that pictures a 360° horizontal and 180° view of a point in space (usually obtained by stitching multi-angular images taken from the same point).
performance based engineering	Designing a structure considering changes during its life-time (usually evaluated probabilistically).
perturbed stress volume	A part of a body whose stresses are no longer balanced.
piecewise linear	Complex function approximated by several linear segments.
pinching	Hysteresis loops that are thinner towards the origin, i.e., show a pronounced change in stiffness after yielding.
portfolio	A (large) set of assets.

Continued on next page

Table G.1 – continued from previous page

Term	Description
premium	Money that is payed to an insurance company (usually annually) for which the insurance guarantees to cover potential losses within the insured period.
principal components	Axis along which a variable has the k-th strongest variance, here only two are used, i.e., strongest and second strongest. The length is defined by the outermost points on these axes.
Quaternary	The geological time period of the Earth (covering the last 2.6 million years).
rake	Angle of relative movement between both sides of the rock along the rupture plane (defined with respect to the longitudinal axis of the rupture plane).
pseudo random	Random number generated on a computer using a seed number that yields a reproducible and thus pseudo random number.
rebar	Steel reinforcement usually in a rod/bar shape.
tectonic regime	Earthquakes are ascribed to different tectonic regimes based on their place of occurrence. One distinguishes (at least) between crustal, interface and intraslab events.
regression	Determining the type of relationship between two (or more) variables.
reinforcement	Steel reinforcement used to compensate for the fact that stone and concrete has little tensile strength.
rifting	Spreading between both sides of a fault.
routable network	A network where each point can be reached from any other point in the network (no isolated points, no dead ends).
rubble	Small stones or parts of stones of various sizes.
rupture kinematics	The dynamic process describing the formation of a rupture plane during an earthquake.
sampling	Selecting individuals from a population.
screening	Analyzing an object visually.
segmentation	Dividing a space into homogeneous (in terms of their attributes) segments.
seismic noise	Ambient vibrations of the ground induced by many natural and man-made sources.

Continued on next page

Table G.1 – continued from previous page

Term	Description
seismic tomography	Imaging the subsurface using several surface recordings of earthquakes whose waves have traveled through the points of interest before being recorded.
semantic	Meaning of something.
slenderness	Ratio between effective column length (here set to storey height, commonly much less) and radius of gyration see Bond et al. (2006, pp. 36,37).
spectral bands	Optical sensors operate with several sensor arrays that capture a signal within an interval of the spectrum of light (a spectral band).
stiffness degradation	Decrease of the stiffness of a material due to permanent damage.
stirrup	Lateral reinforcement embracing the longitudinal reinforcement.
stocky column	Column whose slenderness ratio makes sudden sideways failure (buckling) of the column unlikely.
storey	Structural level of a building at a height level, sometimes referred to as a building floor, but not to be confused with the actual floor of a building.
strain	Change of length in an object with respect to its original length.
strata	Homogeneous subgroup of a population.
sub-plate	Part of a larger tectonic plate that is not fully connected to the rest of the plate.
supervised classification	Classification which is supervised by a person, where usually class labels are assigned manually at some point of the process to elements of the dataset that is being analyzed.
tension	Outwards oriented forces acting on a body.
Tertiary	Geological time period of the Earth ~ 66 till 2.6 million years before today.
textural feature	Visual variation in the pattern of an object.
torsional	Deformation along a rotational axis.
transverse reinforcement	Lateral reinforcement.
tsunami	Large wave in the ocean caused by a strong component of vertical motion during an marine earthquake.
yielding	The amount of deformation from which a material will begin to experience plastic deformation.

H. Software acknowledgment

A variety of software developed by other workers was used in this thesis. The original implementation of the ETAS code was provided by Sebastian Hainzl (GFZ). The first implementation of the RRVS tool was written by Marc Wieland (GFZ). The satellite image analysis used the OTB libraries (OTB Development Team 2014). For the development of the hazard components, parts of the OpenQuake Hazard Modellers ToolKit (Weatherill 2014, HMTK) were used. Ground motion fields, hazard calculations, deaggregation and macroseismic intensity-based risk calculations were performed using OpenQuake (Pagani et al. 2014). Dynamic response was simulated using OpenSees (McKenna 1997). Most programming and scripting was done in Python (van Rossum 1995) and R (Team 2017). For data processing and statistics, mainly pandas (McKinney 2010) and scipy (Jones et al. 2001) were used. Plotting was done using GMT (Wessel and Smith 1991), ggplot2 (Wickham 2009), matplotlib (Hunter 2007) and Qgis (QGIS Development Team 2017). For the computer generated graphics Blender (Blender Online Community 2018) was used. Data storage of exposure data was handled using PostgreSQL (PG 2016) and PostGIS (Ramsey 2005).

Bibliography

- Abrahamson, N A, W J Silva, and R Kamai (2014). “Summary of the ASK14 Ground Motion Relation for Active Crustal Regions”. In: *Earthquake Spectra* 30.3, pp. 1025–1055. ISSN: 8755-2930. DOI: 10.1193/070913EQS198M.
- Abrahamson, N A and R R Youngs (1992). “A Stable Algorithm for Regression Analyses Using the Random Effects Model”. In: *Bulletin of the Seismological Society of America* 82.1, pp. 505–510.
- Agnew, D C (2002). “History of Seismology”. In: *International Handbook of Earthquake and Engineering Seismology - Part A*. Vol. 81. International geophysics series. Amsterdam, Boston, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Academic Press for International Association of Seismology and Physics of the Earth’s Interior, pp. 3–11. ISBN: 0-12-440652-1.
- Aki, K (1956). “Some Problems in Statistical Seismology”. In: *Zisin (Journal of the Seismological Society of Japan. 2nd ser.)* 8.4, pp. 205–228. ISSN: 0037-1114, 1883-9029, 2186-599X. DOI: 10.4294/zisin1948.8.4_205.
- Aki, K (1965). “Maximum Likelihood Estimate of b in the Formula $\log N = a - bM$ and Its Confidence Limits”. In: *Bulletin of the Earthquake Research Institute* 43, pp. 237–239.
- Akkar, S et al. (2010). “The Recently Compiled Turkish Strong Motion Database: Preliminary Investigation for Seismological Parameters”. In: *Journal of Seismology* 14.3, pp. 457–479. ISSN: 1383-4649, 1573-157X. DOI: 10.1007/s10950-009-9176-9.
- Akkar, S et al. (2014). “Reference Database for Seismic Ground-Motion in Europe (RESORCE)”. In: *Bulletin of Earthquake Engineering* 12.1, pp. 311–339. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-013-9506-8.
- Allen, T I and D J Wald (2009). “On the Use of High-Resolution Topographic Data as a Proxy for Seismic Site Conditions (VS30)”. In: *Bulletin of the Seismological Society of America* 99.2A, pp. 935–943. ISSN: 0037-1106. DOI: 10.1785/0120080255.
- Allen, Trevor I, David J Wald, and C Bruce Worden (2012). “Intensity Attenuation for Active Crustal Regions”. In: *Journal of seismology* 16.3, pp. 409–433.
- Amorese, D (2007). “Applying a Change-Point Detection Method on Frequency-Magnitude Distributions”. In: *Bulletin of the Seismological Society of America* 97.5, pp. 1742–1749. ISSN: 0037-1106. DOI: 10.1785/0120060181.
- Ansal, A (2004). *Recent Advances in Earthquake Geotechnical Engineering and Microzonation*. OCLC: 611980288. New York: Kluwer. ISBN: 978-1-4020-2528-0.
- Aschheim, M and E Black (1999). “Effects of Prior Earthquake Damage on Response of Simple Stiffness-Degrading Structures”. In: *Earthquake Spectra* 15.1, pp. 1–24. ISSN: 8755-2930. DOI: 10.1193/1.1586026.

- Atkinson, G M and D J Wald (2007). ““Did You Feel It?” Intensity Data: A Surprisingly Good Measure of Earthquake Ground Motion”. In: *Seismological Research Letters* 78.3, pp. 362–368.
- Avni, R et al. (2002). “Erroneous Interpretation of Historical Documents Related to the Epicenter of the 1927 Jericho Earthquake in the Holy Land”. In: *Journal of seismology* 6.4, pp. 469–476.
- Bach, C and S Hainzl (2012). “Improving Empirical Aftershock Modeling Based on Additional Source Information”. In: *Journal of Geophysical Research: Solid Earth* 117.B4, pp. 1–12. ISSN: 01480227. DOI: 10.1029/2011JB008901.
- Bakun, W H (2006). “MMI Attenuation and Historical Earthquakes in the Basin and Range Province of Western North America”. In: *Bulletin of the Seismological Society of America* 96.6, pp. 2206–2220. ISSN: 0037-1106. DOI: 10.1785/0120060045.
- Barakat, S (2003). *Housing Reconstruction after Conflict and Disaster*. OCLC: 54042453. London: Overseas Development Institute. ISBN: 978-0-85003-695-4.
- Båth, M (1964). “Seismology in the Upper Mantle Project”. In: *Tectonophysics* 1.3, pp. 261–271.
- Bazzurro, P and C A Cornell (1999). “Disaggregation of Seismic Hazard”. In: *Bulletin of the Seismological Society of America* 89.2, pp. 501–520.
- Bazzurro, P et al. (2004). “Guidelines for Seismic Assessment of Damaged Buildings”. In: *Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, Canada*.
- Beauval, C, S Hainzl, and F Scherbaum (2006). “The Impact of the Spatial Uniform Distribution of Seismicity on Probabilistic Seismic-Hazard Estimation”. In: *Bulletin of the Seismological Society of America* 96.6, pp. 2465–2471. ISSN: 0037-1106. DOI: 10.1785/0120060073.
- Beer, F P et al. (2011). *Mechanics of Materials*. 6th ed. New York: McGraw-Hill. ISBN: 978-0-07-338028-5.
- Bielak, J et al. (2010). “The ShakeOut Earthquake Scenario: Verification of Three Simulation Sets”. In: *Geophysical Journal International* 180.1, pp. 375–404. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.2009.04417.x.
- Bindi, D et al. (2011). “Intensity Prediction Equations for Central Asia: Intensity Attenuation Model for Central Asia”. In: *Geophysical Journal International* 187.1, pp. 327–337. ISSN: 0956540X. DOI: 10.1111/j.1365-246X.2011.05142.x.
- Blender Online Community (2018). *Blender - a 3D Modelling and Rendering Package*. Blender Institute, Amsterdam.
- Bommer, J J and F Scherbaum (2008). “The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis”. In: *Earthquake Spectra* 24.4, pp. 997–1009. ISSN: 8755-2930. DOI: 10.1193/1.2977755.
- Bond, A et al. (2006). *How to Design Concrete Structures Using Eurocode 2*. OCLC: 754019783. Surrey, UK: The Concrete Centre. ISBN: 978-1-904818-47-2.
- Boser, B E, I M Guyon, and V N Vapnik (1992). “A Training Algorithm for Optimal Margin Classifiers”. In: *Proceedings of the Fifth Annual ACM Conference on Computational Learning Theory, COLT 1992*. Pittsburgh, PA: ACM Press, pp. 144–152. ISBN: 978-0-89791-497-0. DOI: 10.1145/130385.130401.
- Box, G E P and D R Cox (1964). “An Analysis of Transformations”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 26.2, pp. 211–252. ISSN: 00359246.

- Bozorgnia, Y and V V Bertero (2004). *Earthquake Engineering: From Engineering Seismology to Performance-Based Engineering*. 1st ed. OCLC: 994928429. Boca Raton: CRC Press. ISBN: 978-0-203-48624-5.
- Brown, M B and A B Forsythe (1974). “Robust Tests for the Equality of Variances”. In: *Journal of the American Statistical Association* 69.346, pp. 364–367. ISSN: 0162-1459. DOI: 10.1080/01621459.1974.10482955.
- Brzev, S et al. (2013). *GEM Building Taxonomy Version 2.0*. Tech. rep. GEM Foundation, Pavia, Italy. DOI: 10.13117/GEM.EXP-MOD.TR2013.02.
- Burkhard, M and G Grünthal (2009). “Seismic Source Zone Characterization for the Seismic Hazard Assessment Project PEGASOS by the Expert Group 2 (EG1b)”. In: *Swiss Journal of Geosciences* 102.1, pp. 149–188.
- Çağnan, Z and D Kalafat (2012). “A New Catalogue of Eastern Mediterranean Earthquakes, 2150 BC–2011”. In: *15th World Conference on Earthquake Engineering, Lisboa, Portugal*.
- Campbell, K W and Y Bozorgnia (2008). “NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PGV, PGD and 5% Damped Linear Elastic Response Spectra for Periods Ranging from 0.01 to 10 S ”. In: *Earthquake Spectra* 24.1, pp. 139–171. ISSN: 8755-2930. DOI: 10.1193/1.2857546.
- Campbell, K W and Y Bozorgnia (2014). “NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra”. In: *Earthquake Spectra* 30.3, pp. 1087–1115. ISSN: 8755-2930. DOI: 10.1193/062913EQS175M.
- Cesare, M A et al. (1992). “Modeling Bridge Deterioration with Markov Chains”. In: *Journal of Transportation Engineering* 118.6, pp. 820–833. ISSN: 0733-947X, 1943-5436. DOI: 10.1061/(ASCE)0733-947X(1992)118:6(820).
- Chaljub, E et al. (2010). “Quantitative Comparison of Four Numerical Predictions of 3D Ground Motion in the Grenoble Valley, France”. In: *Bulletin of the Seismological Society of America* 100.4, pp. 1427–1455. ISSN: 0037-1106. DOI: 10.1785/0120090052.
- Chandler, A M and N T K Lam (2002). “Intensity Attenuation Relationship for the South China Region and Comparison with the Component Attenuation Model”. In: *Journal of Asian Earth Sciences* 20.7, pp. 775–790. ISSN: 13679120. DOI: 10.1016/S1367-9120(01)00054-2.
- Chang, C-C and C-J Lin (2011). “LIBSVM: A Library for Support Vector Machines”. In: *ACM transactions on intelligent systems and technology (TIST)* 2.3, p. 27.
- Chen, C-C and Y-X Wu (2006). “An Improved Region-Time-Length Algorithm Applied to the 1999 Chi-Chi, Taiwan Earthquake”. In: *Geophysical Journal International* 166.3, pp. 1144–1148. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.2006.02975.x.
- Cheng, F Y and Y Y Wang (1996). *Post-Earthquake Rehabilitation and Reconstruction*. 1st ed. New York: Pergamon. ISBN: 978-0-08-042825-3.
- Chioccarelli, E and I Iervolino (2015). “Operational Earthquake Loss Forecasting: A Retrospective Analysis of Some Recent Italian Seismic Sequences”. In: *Bulletin of Earthquake Engineering* 14.9, pp. 2607–2626. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-015-9837-8.
- Christophe, E, J Inglada, and A Giros (2008). “Orfeo Toolbox: A Complete Solution for Mapping from High Resolution Satellite Images”. In: *International Archives of*

- the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37.PART B4, pp. 1263–1268.
- Clough, R W and J Penzien (1975). *Dynamics of Structures*. New York: McGraw-Hill. ISBN: 978-0-07-011392-3.
- Cohen, J (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd. Hillsdale, NJ (USA): Laurence Earlbaum Ass.
- Cornell, C A (1968). “Engineering Seismic Risk Analysis”. In: *Bulletin of the Seismological Society of America* 58.5, pp. 1583–1606.
- Crichton, D (1999). “The Risk Triangle”. In: *Natural Disaster Management*. Tudor Rose, pp. 102–103. ISBN: 0-9536140-1-8.
- Crowley, H et al. (2011). *Fragility Functions for Common RC Building Types in Europe*. Deliverable D3.1. Thessaloniki: Aristotle University Thessaloniki, p. 223.
- Al-Dabbeek, J N and A H W Al-Jawhari (2002). *HOUSING REPORT Reinforced Concrete Frame with Infill Walls Designed for Gravity Loading - Palestinian Territories*.
- Al-Dabbeek, J N and R J El-Kelani (2008). “Site Effect and Expected Seismic Performance of Buildings in Palestine- Case Study: Nablus City”. In: *2008 Seismic Engineering Conference Commemorating the 1908 Messina and Reggio Calabria Earthquake*. Ed. by A Santini and N Moraci. American Institute of Physics.
- Dale, R, H Moisl, and H L Somers, eds. (2000). *Handbook of Natural Language Processing*. New York: Marcel Dekker. ISBN: 978-0-8247-9000-4.
- Dijkstra, E W (1959). “A Note on Two Problems in Connexion with Graphs”. In: *Numerische Mathematik* 1, pp. 269–271.
- Douglas, J (2017). *Ground Motion Prediction Equations 1964-2017*. Tech. rep. Glasgow: University of Strathclyde, p. 596.
- Dowrick, D J and D A Rhoades (2005). “Revised Models for Attenuation of Modified Mercalli Intensity in New Zealand Earthquakes”. In: *Bulletin of the New Zealand Society for Earthquake Engineering* 38.4, pp. 185–214.
- Dunford, M and L Li (2011). “Earthquake Reconstruction in Wenchuan: Assessing the State Overall Plan and Addressing the ‘Forgotten Phase’”. In: *Applied Geography* 31.3, pp. 998–1009. ISSN: 01436228. DOI: 10.1016/j.apgeog.2011.01.001.
- Dynes, R R and E L Quarantelli (1989). “Reconstruction in the Context of Recovery: Thoughts on the Alaskan Earthquake”. In: *Preliminary Paper #141*. Vol. 1. Delaware: University of Delaware Disaster Research Center, pp. 1–17.
- Dziewonski, A M, T-A Chou, and J H Woodhouse (1981). “Determination of Earthquake Source Parameters from Waveform Data for Studies of Global and Regional Seismicity”. In: *Journal of Geophysical Research: Solid Earth* 86.B4, pp. 2825–2852. ISSN: 01480227. DOI: 10.1029/JB086iB04p02825.
- Efron, B (1979). “Bootstrap Methods: Another Look at the Jackknife”. In: *Annals of Statistics* 7.1, pp. 1–26.
- Ekström, G, M Nettles, and A M Dziewoński (2012). “The Global CMT Project 2004–2010: Centroid-Moment Tensors for 13,017 Earthquakes”. In: *Physics of the Earth and Planetary Interiors* 200-201, pp. 1–9. ISSN: 00319201. DOI: 10.1016/j.pepi.2012.04.002.
- European Committee for Standardization (2001). *Eurocode 1: Actions on Structures - Part 1-1: General Actions - Densities, Self-Weight, Imposed Loads for Buildings*.
- European Committee for Standardization (2004a). *Eurocode 2: Design of Concrete Structures - Part 1: General Rules and Rules for Buildings*.

- European Committee for Standardization (2004b). *Eurocode 8: Design of Structures for Earthquake Resistance - Part3: Assessment and Retrofitting of Buildings*.
- Faidi, H (2007). “Urban Seismic Risk: ”Impacts on Land Use Planning In Nablus City, Palestine””. PhD thesis. Nablus, Palestine: An-Najah National University.
- Federal Emergency Management Agency (2011). *Hazus–MH 2.1 Multi-Hazard Loss Estimation Methodology - Earthquake Model - Technical Manual*. Tech. rep. Washington, DC: FEMA - Mitigation Division.
- Federal Emergency Management Agency and Rutherford & Chekene (2006). *Techniques for the Seismic Rehabilitation of Existing Buildings*. FEMA 547 - October 2006. OCLC: 946793013. FEMA, nehrp. ISBN: 978-1-78266-134-4.
- Feinerer, I, K Hornik, and D Meyer (2008). “Text Mining Infrastructure in R”. In: *Journal of Statistical Software* 25.5, pp. 1–54. ISSN: 1548-7660. DOI: 10.18637/jss.v025.i05.
- Fekete, A, M Damm, and J Birkmann (2010). “Scales as a Challenge for Vulnerability Assessment”. In: *Natural Hazards* 55.3, pp. 729–747. ISSN: 0921-030X, 1573-0840. DOI: 10.1007/s11069-009-9445-5.
- Felzer, K R and E E Brodsky (2006). “Decay of Aftershock Density with Distance Indicates Triggering by Dynamic Stress”. In: *Nature* 441.7094, pp. 735–738. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature04799.
- Frankel, A (1995). “Mapping Seismic Hazard in the Central and Eastern United States”. In: *Seismological Research Letters* 66.4, pp. 8–21.
- Gardner, J K and L Knopoff (1974). “Is the Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian”. In: *Bull. Seismol. Soc. Am* 64.5, pp. 1363–1367.
- Garfunkel, Z, I Zak, and R Freund (1981). “Active Faulting in the Dead Sea Rift”. In: *Tectonophysics* 80.1, pp. 1–26.
- Gasperini, P et al. (2010). “The Location and Sizing of Historical Earthquakes Using the Attenuation of Macroseismic Intensity with Distance”. In: *Bulletin of the Seismological Society of America* 100.5A, pp. 2035–2066. ISSN: 0037-1106. DOI: 10.1785/0120090330.
- Giovinazzi, S and S Lagomarsino (2004). “A Macroseismic Method for the Vulnerability Assessment of Buildings”. In: *13th World Conference on Earthquake Engineering, Vancouver, BC, Canada*, pp. 1–6.
- Gower, J C (1971). “A General Coefficient of Similarity and Some of Its Properties”. In: *Biometrics* 27.4, pp. 857–871.
- Grigoratos, I et al. (2016). “Development of A Fragility and Exposure Model for Palestine – Application to The City of Nablus”. In: *Procedia Engineering* 161, pp. 2023–2029. ISSN: 18777058. DOI: 10.1016/j.proeng.2016.08.797.
- Grünthal, G and R Wahlström (2012). “The European-Mediterranean Earthquake Catalogue (EMEC) for the Last Millennium”. In: *Journal of Seismology* 16.3, pp. 535–570. ISSN: 1383-4649, 1573-157X. DOI: 10.1007/s10950-012-9302-y.
- Grünthal, G, R Wahlström, and D Stromeyer (2013). “The SHARE European Earthquake Catalogue (SHEEC) for the Time Period 1900–2006 and Its Comparison to the European-Mediterranean Earthquake Catalogue (EMEC)”. In: *Journal of Seismology* 17.4, pp. 1339–1344. ISSN: 1383-4649, 1573-157X. DOI: 10.1007/s10950-013-9379-y.
- Grünthal, G et al. (1998). *European Macroseismic Scale 1998 (EMS-98)*.

- Gu, C et al. (2013). “Triggering Cascades and Statistical Properties of Aftershocks: STATISTICAL PROPERTIES OF AFTERSHOCKS”. In: *Journal of Geophysical Research: Solid Earth* 118.8, pp. 4278–4295. ISSN: 21699313. DOI: 10.1002/jgrb.50306.
- Gutenberg, B and C F Richter (1944). “Frequency of Earthquakes in California”. In: *Bull. Seismol. Soc. Am* 34.4, pp. 185–188.
- Haas, J E, R W. Kates, and M J Bowden (1977). *Reconstruction Following Disaster*. MIT Press environmental studies series. Cambridge, Mass: MIT Press. ISBN: 978-0-262-08094-1.
- Haas, M, J Al-Dabbeek, and M Pittore (2016a). “Constraining Seismic Risk Models by Integrating Remote Sensing and Remote Rapid Visual Screening (RRVS): The DESERVE Experience”. In: *35th General Assembly of the European Seismological Commission*. Trieste, Italy: European Seismological Commission, p. 1.
- Haas, M et al. (2016b). “Data-Driven Seismic-Hazard Models Prepared for a Seismic Risk Assessment in the Dead Sea Region”. In: *Bulletin of the Seismological Society of America* 106.6, pp. 2584–2598. ISSN: 0037-1106, 1943-3573. DOI: 10.1785/0120150361.
- Haas, M et al. (2016c). *DESERVE Earthquake Catalogue and Macroseismic Intensity Dataset*.
- Haas, M et al. (2016d). “Effects of Exposure Model Resolution on Seismic Risk Estimates - Examples from the Cities of Kerak and Madaba in Jordan”. In: *Geophysical Research Abstracts*. Vol. 18. EGU2016-14607, 2016. Vienna.
- Hainzl, S, J Moradpour, and J Davidsen (2014). “Static Stress Triggering Explains the Empirical Aftershock Distance Decay”. In: *Geophysical Research Letters* 41.24, pp. 8818–8824. ISSN: 00948276. DOI: 10.1002/2014GL061975.
- Hancilar, Ufuk and Fabio Taucer, eds. (2013). *Guidelines for Typology Definition of European Physical Assets for Earthquake Risk Assessment*. JRC Scientific and Policy Reports SYNER-G Reference Report 2. Luxembourg: European Commission - Joint Research Centre - Institute for the Protection and Security of the Citizen. ISBN: 978-92-79-28973-6.
- Hanks, T C and H Kanamori (1979). “A Moment Magnitude Scale”. In: *Journal of Geophysical Research* 84.B5, p. 2348. ISSN: 0148-0227. DOI: 10.1029/JB084iB05p02348.
- Hashemi-Parast, S O, F Yamazaki, and W Liu (2017). “Monitoring and Evaluation of the Urban Reconstruction Process in Bam, Iran, after the 2003 M w 6.6 Earthquake”. In: *Natural Hazards* 85.1, pp. 197–213. ISSN: 0921-030X, 1573-0840. DOI: 10.1007/s11069-016-2573-9.
- Helmstetter, A and D Sornette (2002). “Diffusion of Epicenters of Earthquake Aftershocks, Omori’s Law, and Generalized Continuous-Time Random Walk Models”. In: *Physical Review E* 66.6, p. 061104.
- Herrmann, R B (1977). “Recurrence Relations”. In: *Seismological Research Letters* 48.1-2, pp. 47–49. ISSN: 0895-0695, 1938-2057. DOI: 10.1785/gssr1.48.1-2.47.
- Hirano, R (1924). “An Investigation in After Shocks Felt at Kumagai Accompanying the Great Kwanto Earthquake of September 1st, 1923.” In: *Journal of Meteorological Society of Japan* II.2, pp. 77–83.
- Hirayama, Y (2000). “Collapse and Reconstruction: Housing Recovery Policy in Kobe after the Hanshin Great Earthquake”. In: *Housing Studies* 15.1, pp. 111–128. ISSN: 0267-3037, 1466-1810. DOI: 10.1080/02673030082504.

- Hofstetter, A (2003). “Seismic Observations of the 22/11/1995 Gulf of Aqaba Earthquake Sequence”. In: *Tectonophysics* 369.1, pp. 21–36.
- Hofstetter, A et al. (2000). “Crustal and Upper Mantle Structure across the Dead Sea Rift and Israel from Teleseismic P-Wave Tomography and Gravity Data”. In: *Tectonophysics* 327.1-2, pp. 37–59. ISSN: 00401951. DOI: 10.1016/S0040-1951(00)00161-X.
- Hognestad, E (1951). *Study of Combined Bending and Axial Load in Reinforced Concrete Members*. Bulletin Series 399. University of Illinois at Urbana Champaign, College of Engineering. Engineering Experiment Station.
- Holschneider, M, G Zoller, and S Hainzl (2011). “Estimation of the Maximum Possible Magnitude in the Framework of a Doubly Truncated Gutenberg-Richter Model”. In: *Bulletin of the Seismological Society of America* 101.4, pp. 1649–1659. ISSN: 0037-1106. DOI: 10.1785/0120100289.
- Hsu, C-W, C-C Chang, and C-J Lin (2003). *A Practical Guide to Support Vector Classification*.
- Hunter, J D (2007). “Matplotlib: A 2D Graphics Environment”. In: *Computing in Science & Engineering* 9.3, pp. 90–95. ISSN: 1521-9615. DOI: 10.1109/MCSE.2007.55.
- Iervolino, I, M Giorgio, and E Chioccarelli (2016). “Markovian Modeling of Seismic Damage Accumulation: Markovian Modeling of Seismic Damage Accumulation”. In: *Earthquake Engineering & Structural Dynamics* 45.3, pp. 441–461. ISSN: 00988847. DOI: 10.1002/eqe.2668.
- Iervolino, I, M Giorgio, and B Polidoro (2014). “Sequence-Based Probabilistic Seismic Hazard Analysis”. In: *Bulletin of the Seismological Society of America* 104.2, pp. 1006–1012. ISSN: 0037-1106. DOI: 10.1785/0120130207.
- Iervolino, I et al. (2011). “Engineering Ground Motion Record Selection in the Italian Accelerometric Archive”. In: *Bulletin of Earthquake Engineering* 9.6, pp. 1761–1778. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-011-9300-4.
- Jaiswal, K S and D J Wald (2008). *Creating a Global Building Inventory for Earthquake Loss Assessment and Risk Management*. Tech. rep. 1160. U.S. Geological Survey, p. 103.
- Jardaneh, I G (2004). “Evaluation of Ground Response Due to Earthquakes-Case Study”. In: *Journal of Applied Sciences* 4.3, pp. 364–368. ISSN: 18125654, 18125662. DOI: 10.3923/jas.2004.364.368.
- Jardaneh, I G (2007). “Geotechnical Map for the City of Nablus - Palestine”. In: *Journal of An-Najah National University*. Natural Sciences 21, pp. 653–673.
- Jayaram, N and J W Baker (2009). “Correlation Model for Spatially Distributed Ground-Motion Intensities”. In: *Earthquake Engineering & Structural Dynamics* 38.15, pp. 1687–1708. ISSN: 00988847, 10969845. DOI: 10.1002/eqe.922.
- Jiménez-Huerta, D (2009). “Stochastic Models and Methods for the Assessment of Earthquake Risk in Insurance”. PhD thesis. London School of Economics and Political Science.
- Joachims, T (2002). *Learning to Classify Text Using Support Vector Machines*. OCLC: 968635685. Springer-Verlag New York Inc. ISBN: 978-1-4615-0907-3.
- Johnson, K M (2013). “Slip Rates and Off-Fault Deformation in Southern California Inferred from GPS Data and Models”. In: *Journal of Geophysical Research: Solid Earth* 118.10, pp. 5643–5664. ISSN: 21699313. DOI: 10.1002/jgrb.50365.

- Johnston, A C (1996). “Seismic Moment Assessment of Earthquakes in Stable Continental Regions-I. Instrumental Seismicity”. In: *Geophysical Journal International* 124.2, pp. 381–414. ISSN: 0956540X, 1365246X. DOI: 10.1111/j.1365-246X.1996.tb07028.x.
- Jones, E, T Oliphant, and P Peterson (2001). *SciPy: Open Source Scientific Tools for Python*.
- Kagan, Y Y and L Knopoff (1987). “Statistical Short-Term Earthquake Prediction”. In: *Science* 236.4808, pp. 1563–1567. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.236.4808.1563.
- Kale, O and S Akkar (2013). “A New Procedure for Selecting and Ranking Ground-Motion Prediction Equations (GMPEs): The Euclidean Distance-Based Ranking (EDR) Method”. In: *Bulletin of the Seismological Society of America* 103.2A, pp. 1069–1084. ISSN: 0037-1106. DOI: 10.1785/0120120134.
- Kanji, G K (2006). *100 Statistical Tests*. 3rd. London: SAGE Publications Ltd.
- Kaufman, L and P J Rousseeuw (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley series in probability and mathematical statistics. Hoboken, N.J: Wiley. ISBN: 978-0-471-73578-6.
- El-Kelani, R J and I G Jardaneh (2011). “Surface Soil Effects Study Using Microtremor Observations in Nablus City, Palestine”. In: *The Islamic University Journal (Series of Natural Studies and Engineering)* 19.2. OCLC: 56801040, pp. 99–112.
- Kijko, A (2004). “Estimation of the Maximum Earthquake Magnitude, m_{Max} ”. In: *Pure and Applied Geophysics* 161.8, pp. 1655–1681. ISSN: 0033-4553, 1420-9136. DOI: 10.1007/s00024-004-2531-4.
- Klar, A et al. (2011). *Spectral acceleration maps for use in SI 413 amendment No. 5*. in Hebrew. Tech. rep. GII Report No 522/599/11 and NBRI Report No. 2012938-1. Geophysical Institute of Israel.
- Kotha, S R, D Bindi, and F Cotton (2016). “Partially Non-Ergodic Region Specific GMPE for Europe and Middle-East”. In: *Bulletin of Earthquake Engineering* 14.4, pp. 1245–1263. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-016-9875-x.
- Kramer (1996). *Geotechnical Earthquake Engineering*.
- Kunitani, S and T Takada (2008). “Probabilistic Assessment of Buildings Damage Considering Aftershocks of Earthquakes”. In: *13th World Conference on Earthquake Engineering, Vancouver, BC, Canada*.
- Lagomarsino, S and S Giovinazzi (2006). “Macroseismic and Mechanical Models for the Vulnerability and Damage Assessment of Current Buildings”. In: *Bulletin of Earthquake Engineering* 4.4, pp. 415–443. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-006-9024-z.
- Le Goff, B, J F Borges, and M Bezzeghoud (2014). “Intensity-Distance Attenuation Laws for the Portugal Mainland Using Intensity Data Points”. In: *Geophysical Journal International* 199.2, pp. 1278–1285. ISSN: 0956-540X, 1365-246X. DOI: 10.1093/gji/ggu317.
- Lee, D J et al. (2000). *Manual for the Design of Reinforced Concrete Building Structures to EC2*. London: The Institution of Structural Engineers and The Institution of Civil Engineers.
- Levi, T et al. (2015). “Insights from Hazus Loss Estimations in Israel for Dead Sea Transform Earthquakes”. In: *Natural Hazards* 75.1, pp. 365–388. ISSN: 0921-030X, 1573-0840. DOI: 10.1007/s11069-014-1325-y.

- Liel, A B et al. (2013). “Perceptions of Decision-Making Roles and Priorities That Affect Rebuilding after Disaster: The Example of L’Aquila, Italy”. In: *Earthquake Spectra* 29.3, pp. 843–868.
- Luco, N, P Bazzurro, and C A Cornell (2004). “Dynamic versus Static Computation of the Residual Capacity of a Mainshock-Damaged Building to Withstand an Aftershock”. In: *Proceedings of the 13th World Conference on Earthquake Engineering*. Vol. 2405. Vancouver, Canada.
- Luzi, L et al. (2015). *Engineering Strong Motion Database*. DOI: 10.13127/ESM.
- Maffei, J et al. (2004). *Test Applications of Advanced Seismic Assessment Guidelines*. Tech. rep. Oakland, California: RUTHERFORD & CHEKENE Consulting Engineers, p. 77.
- Mak, S, R A Clements, and D Schorlemmer (2014). “The Statistical Power of Testing Probabilistic Seismic-Hazard Assessments”. In: *Seismological Research Letters* 85.4, pp. 781–783. ISSN: 0895-0695, 1938-2057. DOI: 10.1785/0220140012.
- Mann, H B and D R Whitney (1947). “On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other”. In: *The Annals of Mathematical Statistics* 18.1, pp. 50–60. ISSN: 0003-4851. DOI: 10.1214/aoms/1177730491.
- Massey, F J (1951). “The Kolmogorov-Smirnov Test for Goodness of Fit”. In: *Journal of the American Statistical Association* 46.253, p. 68. ISSN: 01621459. DOI: 10.2307/2280095.
- Masson, F et al. (2015). “Variable Behavior of the Dead Sea Fault along the Southern Arava Segment from GPS Measurements”. In: *Comptes Rendus Geoscience* 347.4, pp. 161–169. ISSN: 16310713. DOI: 10.1016/j.crte.2014.11.001.
- McKenna, F (1997). “Object-Oriented Finite Element Programming Frameworks for Analysis, Algorithms and Parallel Computing”. Dissertation. Berkley, United States: University of California, Berkley.
- McKinney, W (2010). “Data Structures for Statistical Computing in Python”. In: *Proceedings of the 9th Python in Science Conference (SciPy2010)*. Austin, Texas, pp. 51–56.
- Metropolis, N and S Ulam (1949). “The Monte Carlo Method”. In: *Journal of the American Statistical Association* 44.247, p. 335. ISSN: 01621459. DOI: 10.2307/2280232.
- Milutinovic, Z V and G S Trendafiloski (2003). *WP4 Vulnerability of Current Buildings*. Tech. rep. European Commission, p. 110.
- Moradpour, J, S Hainzl, and J Davidsen (2014). “Nontrivial Decay of Aftershock Density with Distance in Southern California”. In: *Journal of Geophysical Research: Solid Earth* 119.7, pp. 5518–5535. ISSN: 21699313. DOI: 10.1002/2014JB010940.
- Mucciarelli, M and M R Gallipoli (2006). “Comparison between Vs30 and Other Estimates of Site Amplification in Italy”. In: *First European Conference on Earthquake Engineering and Seismology (a joint event of the 13th ECEE & 30th General Assembly of the ESC), Geneva Switzerland, 3-8 Spetember 2006, Paper Number 270*.
- Nakamura, Y (1989). “A Method for Dynamic Characteristics Estimation of Subsurface Using Microtremor on the Ground Surface”. In: *Quarterly Report of the Railway Technical Research Institute* 30.1, pp. 25–33.
- Ogata, Y (1988). “Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes”. In: *Journal of the American Statistical Associa-*

- tion 83.401, pp. 9–27. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1988.10478560.
- Omori, F (1894). “On After-Shocks of Earthquakes”. In: *The journal of the College of Science, Imperial University, Japan* 7, pp. 111–200.
- OpenStreetMap contributors (2017). *Planet Dump Retrieved from <https://planet.osm.org>*.
- OTB Development Team (2014). *The ORFEO Tool Box Software Guide Updated for OTB-4.2*. Tech. rep. Centre National D’Études Spatiales cnes.
- Pagani, M et al. (2014). “OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model”. In: *Seismological Research Letters* 85.3, pp. 692–702. ISSN: 0895-0695, 1938-2057. DOI: 10.1785/0220130087.
- Pantazopoulou, S J and D V Syntzirma (2010). “Deformation Capacity of Lightly Reinforced Concrete Members - Comparative Evaluation”. In: *Advances in Performance-Based Earthquake Engineering: ACES Workshop*. Dordrecht ; New York: Springer, p. 486. ISBN: 978-90-481-8745-4.
- Paolucci, R et al. (2014). “Physics-Based Earthquake Ground Shaking Scenarios in Large Urban Areas”. In: *Perspectives on European Earthquake Engineering and Seismology*. Ed. by Atilla Ansal. Vol. 34. Cham: Springer International Publishing, pp. 331–359. ISBN: 978-3-319-07117-6. DOI: 10.1007/978-3-319-07118-3_10.
- Papadopoulos, A and P Bazzurro (2018). “Sensitivity Analysis of Earthquake Loss Estimation Using the Space-Time ETAS Model for Seismicity Clustering”. In: *16th European Conference on Earthquake Engineering 18-21 June 2018*. Thessaloniki: European Association of Earthquake Engineering, pp. 1–12.
- Parolai, S (2012). “Investigation of Site Response in Urban Areas by Using Earthquake Data and Seismic Noise”. In: *New manual of seismological observatory practice* 2, pp. 1–38.
- Parzen, E (2015). *Stochastic Processes*. Dover edition. Mineola, New York: Dover Publications, Inc. ISBN: 978-0-486-79688-8.
- Pasolini, C et al. (2008). “The Attenuation of Seismic Intensity in Italy, Part I: Theoretical and Empirical Backgrounds”. In: *Bulletin of the Seismological Society of America* 98.2, pp. 682–691.
- PG, PostgreSQL Global Development Group (2016). *PostgreSQL*.
- pgRouting Contributors (2013). *pgRouting Manual Release 2.0.0*.
- Pinto, A et al. (2002). *Pseudo-Dynamic Tests on Non-Seismic Resisting RC Frames (Bare and Selective Retrofit Frames)*. Tech. rep. EUR 20244 EN. Ispra, Italy: European Commission - Joint Research Centre -Institute for the Protection and the Security of the Citizen European Laboratory for Structural Assessment (ELSA), p. 175.
- Pittore, M, M Haas, and K Megalooikonomou (2018). “Risk-Oriented , Bottom-up Modelling of Building Portfolios with Faceted Taxonomies”. In: *Frontiers in Built Environment* accepted.
- Pittore, M and M Wieland (2013). “Toward a Rapid Probabilistic Seismic Vulnerability Assessment Using Satellite and Ground-Based Remote Sensing”. In: *Natural hazards* 68.1, pp. 115–145.
- Pittore, M, M Wieland, and K Fleming (2017). “Perspectives on Global Dynamic Exposure Modelling for Geo-Risk Assessment”. In: *Natural Hazards* 86.S1, pp. 7–30. ISSN: 0921-030X, 1573-0840. DOI: 10.1007/s11069-016-2437-3.
- Pittore, M et al. (2015). “Improving Post-Earthquake Insurance Claim Management: A Novel Approach to Prioritize Geospatial Data Collection”. In: *ISPRS*

- International Journal of Geo-Information* 4.4, pp. 2401–2427. ISSN: 2220-9964. DOI: 10.3390/ijgi4042401.
- Priestley, M J N, G M Calvi, and M J Kowalsky (2007). *Displacement-Based Seismic Design of Structures*. OCLC: ocn173622581. Pavia, Italy: IUSS Press : Distributed by Fondazione EUCENTRE. ISBN: 978-88-6198-000-6.
- QGIS Development Team (2017). *QGIS Geographic Information System*.
- Ramsey, P (2005). *PostGIS Manual*.
- RMS (2005). *Hurricane Katrina: Profile of a Super Cat - Lessons and Implications for Catastrophe Risk Management*.
- Rodriguez-Marek, A et al. (2011). “Analysis of Single-Station Standard Deviation Using the KiK-Net Data”. In: *Bulletin of the Seismological Society of America* 101.3, pp. 1242–1258. ISSN: 0037-1106. DOI: 10.1785/0120100252.
- Rojahn, C et al. (1985). *Earthquake Damage Evaluation Data for California*. Redwood City, Calif.: ATC.
- Russell, J E (1994). “Post Earthquake Reconstruction Regulation by Local Government”. In: *Earthquake Spectra* 10.1, pp. 209–223. ISSN: 8755-2930. DOI: 10.1193/1.1585770.
- Salamon, A et al. (1996). “Seismicity of the Eastern Mediterranean Region: Perspective from the Sinai Subplate”. In: *Tectonophysics* 263.1, pp. 293–305.
- Sammon, J W (1969). “A Nonlinear Mapping for Data Structure Analysis”. In: *IEEE Transactions on Computers* C-18.5, pp. 401–409.
- Scherbaum, F, E Delavaud, and C Riggelsen (2009). “Model Selection in Seismic Hazard Analysis: An Information-Theoretic Perspective”. In: *Bulletin of the Seismological Society of America* 99.6, pp. 3234–3247. ISSN: 0037-1106. DOI: 10.1785/0120080347.
- Scherbaum, F et al. (2010). “Exploring the Proximity of Ground-Motion Models Using High-Dimensional Visualization Techniques”. In: *Earthquake Spectra* 26.4, pp. 1117–1138. ISSN: 8755-2930. DOI: 10.1193/1.3478697.
- Schneider, K-J, A Goris, and A Albert, eds. (2012). *Bautabellen für Ingenieure: mit Berechnungshinweisen und Beispielen*. in German. 20th ed. OCLC: 845000221. Köln: Werner. ISBN: 978-3-8041-5251-9.
- Scordilis, E M (2006). “Empirical Global Relations Converting M S and m b to Moment Magnitude”. In: *Journal of Seismology* 10.2, pp. 225–236. ISSN: 1383-4649, 1573-157X. DOI: 10.1007/s10950-006-9012-4.
- SEAOC (1995). *Vision 2000: Performance Based Seismic Engineering of Buildings*. Sacramento, California: Structural Engineers Association of California.
- Şeşetyan, K et al. (2018). “The 2014 Seismic Hazard Model of the Middle East: Overview and Results”. In: *Bulletin of Earthquake Engineering*, pp. 1–32. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-018-0346-4.
- Shearer, P M (2011). *Introduction to Seismology*. 2. ed., repr. with corr. OCLC: 740957394. Cambridge: Cambridge Univ. Press. ISBN: 978-0-521-70842-5.
- Smerzini, C and K Pitilakis (2018). “Seismic Risk Assessment at Urban Scale from 3D Physics-Based Numerical Modeling: The Case of Thessaloniki”. In: *Bulletin of Earthquake Engineering* 16.7, pp. 2609–2631. ISSN: 1570-761X, 1573-1456. DOI: 10.1007/s10518-017-0287-3.
- Sobolev, G and Y S Tyupkin (1997). “Low-Seismicity Precursors of Large Earthquakes in Kamchatka”. In: *Volcanology & Seismology* 18.4 C, pp. 433–445.

- Song, R, Y Li, and J W van de Lindt (2014). “Loss Estimation of a Code-Conforming Steel Building to Mainshock-Aftershock Sequences”. In: *Proceedings of the 10th National Conference in Earthquake Engineering*. Anchorage, Alaska: Earthquake Engineering Research Institute.
- Sørensen, M B, D Stromeyer, and G Grünthal (2009). “Attenuation of Macroseismic Intensity: A New Relation for the Marmara Sea Region, Northwest Turkey”. In: *Bulletin of the Seismological Society of America* 99.2A, pp. 538–553.
- Sørensen, Mathilde B, D Stromeyer, and G Grünthal (2010). “A Macroseismic Intensity Prediction Equation for Intermediate Depth Earthquakes in the Vrancea Region, Romania”. In: *Soil Dynamics and Earthquake Engineering* 30.11, pp. 1268–1278. ISSN: 02677261. DOI: 10.1016/j.soildyn.2010.05.009.
- Steinhaus, H (1956). “Sur La Division Des Corps Matériels En Parties”. In: *Bulletin de l'Académie Polonaise des Sciences* CL. III -Vol. IV.12, pp. 801–804.
- Stephens, M A (1974). “EDF Statistics for Goodness of Fit and Some Comparisons”. In: *Journal of the American Statistical Association* 69.347, pp. 730–737. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1974.10480196.
- Stepp, J C (1971). “An Investigation of Earthquake Risk in the Puget Sound Area by Use of the Type I Distribution of Largest Extremes”. PhD Thesis. Pennsylvania: Pennsylvania State University.
- Stucchi, M et al. (2013). “The SHARE European Earthquake Catalogue (SHEEC) 1000–1899”. In: *Journal of Seismology* 17.2, pp. 523–544. ISSN: 1383-4649, 1573-157X. DOI: 10.1007/s10950-012-9335-2.
- Team, R Core (2017). *R: A Language and Environment for Statistical Computing*. Vienna.
- Tesfamariam, S and K Goda (2015). “Loss Estimation for Non-Ductile Reinforced Concrete Building in Victoria, British Columbia, Canada: Effects of Mega-Thrust M_w 9-Class Subduction Earthquakes and Aftershocks: Loss Estimation for Non-Ductile Reinforced Concrete Building”. In: *Earthquake Engineering & Structural Dynamics* 44.13, pp. 2303–2320. ISSN: 00988847. DOI: 10.1002/eqe.2585.
- Tesfamariam, S and K Goda (2017). “Impact of Earthquake Types and Aftershocks on Loss Assessment of Non-Code Conforming Buildings: Case Study with Victoria, British Columbia”. In: *Earthquake Spectra* 33.2, pp. 551–579.
- Trevlopoulos, K and P Guéguen (2016). “Period Elongation-Based Framework for Operative Assessment of the Variation of Seismic Vulnerability of Reinforced Concrete Buildings during Aftershock Sequences”. In: *Soil Dynamics and Earthquake Engineering* 84, pp. 224–237. ISSN: 02677261. DOI: 10.1016/j.soildyn.2016.02.009.
- Ulam, S, R D Richtmyer, and J von Neumann (1947). *Statistical Methods in Neutron Diffusion*. LAMS-551. Los Alamos Scientific Laborator.
- Utsu, T (1965). “A Method for Determining the Value of b in a Formula $\log N = a - bM$ Showing the Magnitude-Frequency Relation for Earthquakes”. In: *Geophysical Bulletin of Hokkaido University* 13, pp. 99–103.
- Utsu, T and Y Ogata (1995). “The Centenary of the Omori Formula for a Decay Law of Aftershock Activity.” In: *Journal of Physics of the Earth* 43.1, pp. 1–33.
- Vamvatsikos, D and A C Cornell (2006). “Direct Estimation of the Seismic Demand and Capacity of Oscillators with Multi-Linear Static Pushovers through IDA”. In: *Earthquake Engineering & Structural Dynamics* 35.9, pp. 1097–1117. ISSN: 0098-8847, 1096-9845. DOI: 10.1002/eqe.573.

- Vamvatsikos, D and C A Cornell (2002). “Incremental Dynamic Analysis”. In: *Earthquake Engineering & Structural Dynamics* 31.3, pp. 491–514.
- Vamvatsikos, D and S J Pantazopoulou (2016). “Simplified Mechanical Model to Estimate the Seismic Vulnerability of Heritage Unreinforced Masonry Buildings”. In: *Journal of Earthquake Engineering* 20.2, pp. 298–325. ISSN: 1363-2469, 1559-808X. DOI: 10.1080/13632469.2015.1060583.
- van Rossum, G (1995). *Python Library Reference*. Tech. rep. CS-R9524 1995. Amsterdam: CWI - Centrum voor Wiskunde en Informatica, Stichting Mathematisch Centrum, p. 173.
- van Stiphout, T, J Zhuang, and D Marsan (2012). “Seismicity Declustering”. In: *CORSSA v.1.0*. DOI: 10.5078/corssa-52382934.
- Vere-Jones, D (1966). “A Markov Model for Aftershock Occurrence”. In: *pure and applied geophysics* 64.1, pp. 31–42.
- Vidic, T, P Fajfar, and M Fischinger (1994). “Consistent Inelastic Design Spectra: Strength and Displacement”. In: *Earthquake Engineering & Structural Dynamics* 23.5, pp. 507–521. ISSN: 00988847, 10969845. DOI: 10.1002/eqe.4290230504.
- Waddell, P (2002). “UrbanSim: Modeling Urban Development for Land Use, Transportation, and Environmental Planning”. In: *Journal of the American Planning Association* 68.3, pp. 297–314. ISSN: 0194-4363, 1939-0130. DOI: 10.1080/01944360208976274.
- Wald, D J and T I Allen (2007). “Topographic Slope as a Proxy for Seismic Site Conditions and Amplification”. In: *Bulletin of the Seismological Society of America* 97.5, pp. 1379–1395. ISSN: 0037-1106. DOI: 10.1785/0120060267.
- Weatherill, G A (2014). *Hazard Modeller’s Toolkit - User Guide*. Technical Report. Global Earthquake Model (GEM).
- Weggel, D C, D M Boyajian, and S Chen (2007). “Modelling Structures as Systems of Springs”. In: *World Transactions on Engineering and Technology Education* 6.1, pp. 169–172.
- Welch, B L (1947). “The Generalization of ‘Student’s’ Problem When Several Different Population Variances Are Involved”. In: *Biometrika* 34.1/2, p. 28. ISSN: 00063444. DOI: 10.2307/2332510.
- Wells, D L and K J Coppersmith (1994). “New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement”. In: *Bulletin of the seismological Society of America* 84.4, pp. 974–1002.
- Wessel, P and W H F Smith (1991). “Free Software Helps Map and Display Data”. In: *Eos Trans. AGU* 72.41, pp. 441–448.
- Wickham, H (2009). *Ggplot2: Elegant Graphics for Data Analysis*. Use R! OCLC: ocn382399721. New York: Springer. ISBN: 978-0-387-98140-6.
- Wieland, M (2013). “Exposure Estimation for Rapid Seismic Vulnerability Assessment: An Integrated Approach Based on Multi-Source Imaging”. PhD Thesis. Berlin: Technische Universität Berlin.
- Wieland, M et al. (2012). “Estimating Building Inventory for Rapid Seismic Vulnerability Assessment: Towards an Integrated Approach Based on Multi-Source Imaging”. In: *Soil Dynamics and Earthquake Engineering* 36, pp. 70–83.
- Wieland, M et al. (2015). “A Multiscale Exposure Model for Seismic Risk Assessment in Central Asia”. In: *Seismological Research Letters* 86.1, pp. 210–222. ISSN: 0895-0695, 1938-2057. DOI: 10.1785/0220140130.

- Wiemer, S (2000). “Introducing Probabilistic Aftershock Hazard Mapping”. In: *Geophysical Research Letters* 27.20, pp. 3405–3408. ISSN: 00948276. DOI: 10.1029/2000GL011479.
- Wilcoxon, F (1945). “Individual Comparisons by Ranking Methods”. In: *Biometrics Bulletin* 1.6, p. 80. ISSN: 00994987. DOI: 10.2307/3001968.
- Woo, G (1996). “Kernel Estimation Methods for Seismic Hazard Area Source Modeling”. In: *Bulletin of the Seismological Society of America* 86.2, pp. 353–362.
- Worden, C B et al. (2012). “Probabilistic Relationships between Ground-Motion Parameters and Modified Mercalli Intensity in California”. In: *Bulletin of the Seismological Society of America* 102.1, pp. 204–221. ISSN: 0037-1106. DOI: 10.1785/0120110156.
- Yeo, G L and C A Cornell (2005). *A Probabilistic Framework for Quantification of Aftershock Ground-Motion Hazard in California: Methodology and Parametric Study*. Tech. rep. 2005/13. Pacific Earthquake Engineering Research Center.
- Yeo, G L and C A Cornell (2009a). “Building Life-Cycle Cost Analysis Due to Mainshock and Aftershock Occurrences”. In: *Structural safety* 31.5, pp. 396–408. ISSN: ISSN 0167-4730.
- Yeo, G L and C A Cornell (2009b). “Post-Quake Decision Analysis Using Dynamic Programming”. In: *Earthquake Engineering & Structural Dynamics* 38.1, pp. 79–93. ISSN: 00988847, 10969845. DOI: 10.1002/eqe.842.
- Yepes-Estrada, C et al. (2017). “Modeling the Residential Building Inventory in South America for Seismic Risk Assessment”. In: *Earthquake Spectra* 33.1, pp. 299–322. ISSN: 8755-2930. DOI: 10.1193/101915EQS155DP.
- Zaslavsky, Y et al. (2012). “Questioning the Applicability of Soil Amplification Factors as Defined by NEHRP (USA) in the Israel Building Standards”. In: *Natural Science* 04.08, pp. 631–639. ISSN: 2150-4091, 2150-4105. DOI: 10.4236/ns.2012.428083.
- Zöller, G and M Holschneider (2016). “The Earthquake History in a Fault Zone Tells Us Almost Nothing about m_{Max} ”. In: *Seismological Research Letters* 87.1, pp. 132–137. ISSN: 0895-0695, 1938-2057. DOI: 10.1785/0220150176.
- Zöller, G et al. (2014). “The Largest Expected Earthquake Magnitudes in Japan: The Statistical Perspective”. In: *Bulletin of the Seismological Society of America* 104.2, pp. 769–779. ISSN: 0037-1106. DOI: 10.1785/0120130103.