HOW TO BUILD AN EARTHQUAKE LOSS MODEL

Lead Author: Patrice Tscherrig – SwissRe & Mohammad Zolfaghari, CatRisk

Table of Contents

HOW TO BUILD AN EARTHQUAKE LOSS MODEL

Lead Author: Patrice Tscherrig – SwissRe & Mohammad Zolfaghari, CatRisk

AUDIENCE

The audience for this paper will be researchers within the seismological, geophysical or earthquake engineering fields who have some knowledge of probabilistic seismic hazard assessment (PSHA) but who are likely to be unfamiliar with the application of PSHA to an insurance loss model.

GOALS AND APPLICATION

Before starting to build an earthquake loss model the practical usage of the model has to be carefully evaluated. In this paper the focus is on an earthquake risk model used to assess property losses for single-locations or portfolios of values due to ground shaking.

If the model is to be used in a different context other elements not mentioned here might become important and, equally, some aspects might become irrelevant. It is even perceivable that an entirely different approach might need to be taken. For instance, if the model is required to reflect an emergency situation in a nuclear power plant the correlation of ground-motions might become irrelevant but the assessment of the specific site conditions much more important.

Potential sources of loss to a single location or to a portfolio of values must be very carefully evaluated. For example, from a property insurance perspective additional loss drivers that occur during an earthquake event, such as earthquake fire-following, might become very relevant. Others, such as tsunami, might become irrelevant if they were excluded from the insurance policies.

In this paper we focus on the losses triggered by only earthquake ground-shaking. Discussion of other important sources of property loss, such as tsunami, fire-following, landslide, liquefaction, fault displacement, seiches and others can be found in similar papers.

This article aims to provide guidelines to how earthquake loss models are created by incorporating scientific knowledge and research expertise in the field of earthquake engineering. The first step towards rational management of earthquake risk is a thorough understanding of severity and frequency distribution of associated risk. Probabilistic catastrophe loss models are ideally suited to fulfil the growing demands of the insurance and reinsurance industries. In the case of earthquake, such models provide probabilistic estimates of seismic losses, defined by loss severity and frequency. Such estimates are achieved through full probabilistic analyses, taking into account all potential earthquake sources and associated uncertainties.

Several factors influence the severity and frequency of building damage caused by earthquake. In general, an insurance-based seismic risk model consists of several main components as shown in Figure (1).

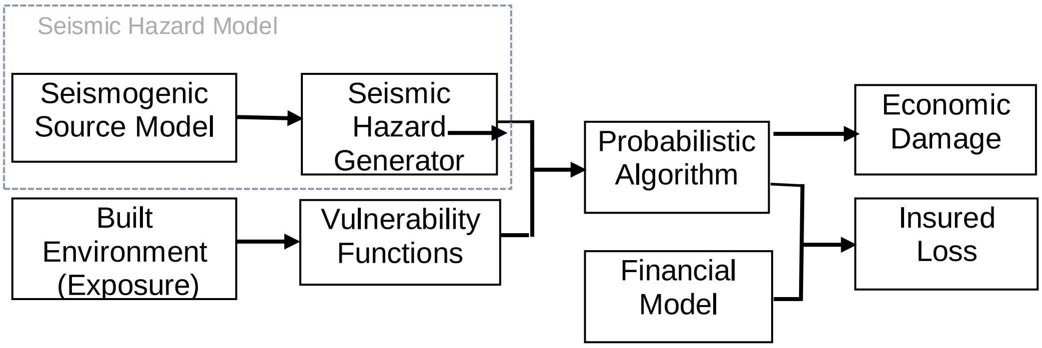

Figure 1: Main components of an insurance-based seismic risk model

SEISMIC HAZARD MODEL

There are several direct and indirect hazards relating to slip along tectonic faults and each can translate into building damage depending on both the hazard type and the built environment characteristics. These hazards include ground shaking, surface rupture, tsunami, landslide, liquefaction, fire-following and water seiches. The severity of the damage related to these hazards during an earthquake depends on many regional factors. The most common and widely spread damage during most earthquakes however is due to strong ground motion, known as ground shaking. This is particularly the case for buildings, which generally form the largest part of insurable assets. Some of the other earthquake-related hazards such as landslide, liquefaction and fire-following are also results of ground shaking but non-shaking hazards are the dominant hazards for lifeline and urban infrastructures.

Geographical distribution of earthquake sources and their induced seismic ground motion can be categorized and modelled by a seismic hazard model. The term “seismic hazard” refers to a probabilistic measure, represented by severity and frequency of ground motion, which is produced from full probabilistic seismic hazard analyses. Different methods may be used for seismic hazard assessment depending on the purpose for which the assessment is required. The method used may depend on the availability of information such as seismicity data, tectonic characteristics and site conditions in addition to the estimation of ground motion force. Seismic hazard assessment provides the base for both pre- and post-catastrophe planning. The most common applications of such an assessment are in seismic design code development and for engineering design purposes. There are other applications however, including disaster mitigation, land and town planning, emergency and post-earthquake response management and not least in insurance catastrophe loss modelling.

Engineers are commonly interested in the severity of ground shaking for return periods (also known as recurrence or repeat intervals – an estimate of the likelihood of a catastrophic event to occur) of between 500 or 1000 years and typically at a single location. In catastrophe loss models however, due to full integration of uncertainties associated with shaking intensity and building vulnerabilities, a full spectrum of shaking hazard from both low to medium *and* high return periods is used. It is important that the correlation of shaking across multiple sites is correct as this is the key driver of how catastrophic losses will translate into insured losses.

There are always uncertainties associated with the development and use of seismic hazard models. While some of these uncertainties can be controlled by more accurate and reliable input data, the majority remain because they are intrinsic to our state of knowledge. One of the main purposes of a catastrophe model is to quantify these uncertainties and to assess their impact on the final result.

There is no single accepted method of building an earthquake loss model. The following sections however outline some of the more important aspects and resources that must be considered, whatever method is used. This guide is not meant to be a construction manual, but rather an overview of existing methods and approaches. It is really down to the end-user - who has the application of the model in mind – to pick the right ingredients and customize it to its end-purpose.

Stochastic event set generation

To assess earthquake hazard the well-established approach of a probabilistic seismic hazard assessment (PSHA) is conducted. In contrast to the typical application of a hazard or design-load map, the ultimate goal is to derive a probability distribution of seismic hazard for the set of assets. As financial conditions can be quite complex, treatment of correlation requires loss assessment methodology which tackle each earthquake scenario separately. Therefore, standard PSHA methodology which relies on numerical integration of total probability equation is not practicable for proper loss calculation. The most promising approach in the context of earthquake loss models is to use Monte Carlo methods (e.g. (Musson, 2000))

The premise is that a catalogue is to be built of possible earthquake losses reflecting a potential loss history to the portfolio (or single location) of interest. This catalogue has to meet the following boundary conditions:

- It has to be consistent with a traditional PSHA conducted in the study region
- It has to be 'small' enough that a portfolio analysis can be computed in a meaningful time with meaningful computer power
- It has to be 'large' enough that it reflects the potential loss outcomes of several 10,000 years – for portfolios *and* for single locations. The analysis has to deliver stable results and cannot be overly sensitive to the effects of random sampling.

Modelling of Earthquake sources

One of the first steps in any seismic hazard study is the definition of seismogenic sources. A source model represents distribution of future earthquakes in space, time and size. Definition of seismic sources could be based on very detailed information with regard to the geometry and seismic potential of causative faults or simply based on some kind of generalization of regional seismicity into basic space and time distributions.

A meaningful approach in producing an event set is to differentiate between *known* seismic sources (faults and subduction zones) and *unknown* seismic sources (background events).

Fault specific source model

Potential seismic activity along a fault rupture can be modelled in simple or more sophisticated ways depending on the availability of fault information as well as its reliability.

Fault models consider the geometry of causative faults. The simplest geometric features used to define rupture geometry in many seismic hazard tools are line segments, each representing the projection of a vertical rupture plane on the surface. Regardless of the simplicity of such models, the fault-to-site distance definition used in ground motion prediction equations (GMPE - discussed further **Error! Reference source not found.**) has to match the source model. To achieve a better calculation of probabilistic ground motion close to a large earthquake, it is necessary to take into account the 3D geometry of fault rupture. This also takes account of other factors influencing near-fault ground motions, such as footwall vs. hanging wall positions. To properly model a fault rupture plane and address source-to-site definitions at near-fault positions, specific geometric information is required to construct a 3D representation of a rupture plane. Figure (2) shows such parameters and how they relate to each other.

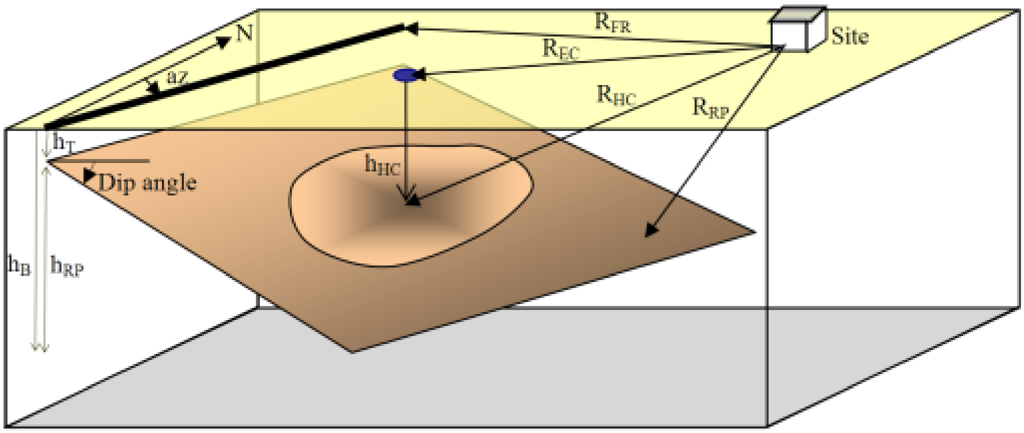

Figure (2): Earthquake fault geometry and various measurements of source-to-site distance used in GMPE.

Scaling relations (e.g. (Wells & Coppersmith, 1994)) can be used to estimate the rupture geometry in case the fault segmentation is not well known. The maximum magnitude can be inferred from the maximum fault area on the segment of interest.

The *source geometry* is also important in the context of loss modelling as it will drive the correlation of losses across multiple sites. For a traditional PSHA it may not be important if a specific ground motion level of interest is caused by a small or big earthquake. However, in respect of the financial loss to a portfolio, this is an important consideration since location correlation and financial constructs such as deductibles will be affected.

After establishing the geometries of the faults the *activity* of the faults must be determined. The rates of the events can be set using *paleoseismology*, which uses historical earthquake catalogues and is based on slip rate measurements.

Logic trees (Bommer & Scherbaum, The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis, 2008) are typically used to capture the uncertainty of fault geometries or fault activities. The design of these logic trees has to be carefully considered.

Area Seismic Source Model

For regions where seismic activity is generated by diffused patterns of small to moderate faults, earthquake sources are modeled by area source zones. These are usually defined by a combination of historical seismicity and characteristics of tectonic features. Area sources are also used in cases where exact fault locations are not known in detail. Earthquakes are assumed to occur anywhere within these sources with a uniform or non-uniform probability distribution (e.g. Zolfaghari, 2015) (Hiemer, et al., 2014). Generalization of regional seismicity into defined area source boundaries involves expert interpretation of seismotectonic information which therefore introduces a subjective element into probabilistic seismic hazard assessment. This is particularly the case for regions with low seismicity and where earthquakes cannot be attributed to specific tectonic features.

When the seismic zones do not fully cover the seismogenic area, background seismicity is often introduced to account for the random occurrence of earthquakes that are not attributable to a known source such as a fault or source zone. Various methods exist to spatially smooth the input catalogue such as Gaussian smoothing using *fixed* filter lengths ((Frankel, 1996)), Gaussian smoothing using *adaptive* filter lengths (e.g. (Field, Biasi, Bird, & et al., 2013)) or kernel functions (Hiemer, et al., 2014).

Subduction Zones

In subduction regions, earthquakes of various sizes, fault mechanism and depth might be located across the zone and related to different layers of tectonic engagement. While the biggest and mega-thrust events are usually associated with a main subduction interface, moderate to big events can also happen at shallow and deep interfaces. Deep earthquakes are related to deformation taking place in the subducting plates at depths of 100 to 200 km, known as Wadati-Benioff earthquakes (Figure 4). In order to simulate potential seismic activity in such areas therefore, several layers of seismic sources need to be modelled to capture the characteristics of seismicity generated at different depths. The subduction interface is in most cases the dominant source of the seismic hazard. Capturing its geometry in the 3D plane is key (e.g. (Hayes, Wald, & Johnson, 2012))

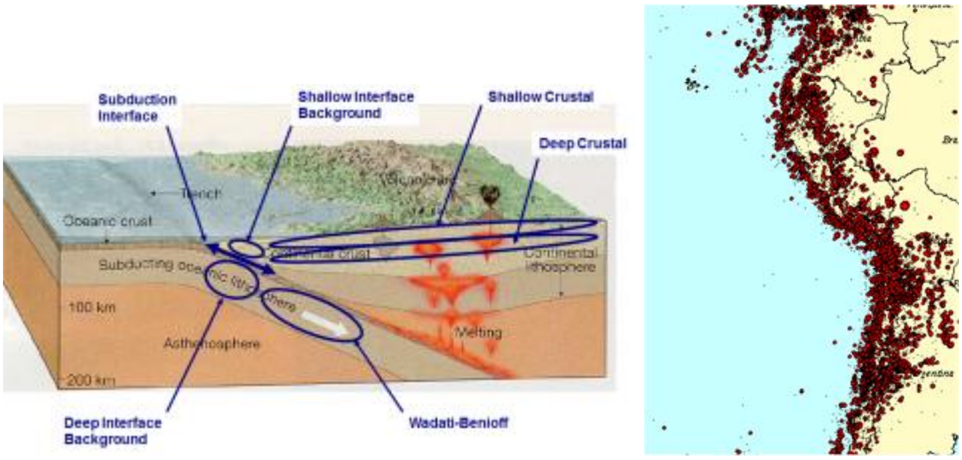

Figure 4: Source of seismicity in areas engaged in subduction tectonic deformation

Because of the significant deformation taking place in some of the major subduction regions and also the size of plate interfaces, assessment of earthquake size and recurrence rate should take account of both slip rate and of locking area along tectonic plates in subduction regions. Historical earthquake data or, where available, paleoseismological information may be used to assess the rate and size of earthquakes in such regions. The maximum magnitude should also be taken into account. (Strasser, Arango, & Bommer, 2010). It is of practical use to apply the logic tree approach to model uncertainties associated with all information on earthquake subduction regions.

Earthquake catalogue

Earthquake catalogues used in seismic hazard assessment usually consists of three main sub-catalogues. For historical earthquakes prior to instrumental recording, the location and magnitude have been estimated based on only macroseismic information collected from multidisciplinary studies of historical sources (e.g Ambraseys and Melville 1982, (Bakun , Johnston , & Hopper , 2003)). The second sub catalogue of earthquake data comes after the development of seismographic instruments of some quantity and quality (about 1900 through to 1963 depending on the region). Seismologists have revisited data from certain countries in order to recalculate the epicentral locations and magnitudes (e.g Ambraseys and Melville, 1996). The third sub catalogue of earthquake catalogue derives from the modern instrumental period and obviously contains the most complete and homogenous data. There are numerous agencies around the world collecting and processing data from global seismological networks (e.g. International Seismological Centre, ISC) and attempts are being made to homogenise these various data sources (Storchak, Giacomo, Bondàr, Engdahl, & et al, 2013).

Seismogenic Parameters for each Source

Gutenberg and Richter frequency-magnitude relationship (for area seismic sources) and characteristic earthquake models (Schwartz & Coppersmith, 1984)for fault-specific sources are the basis of probabilistic seismic hazard analysis. These are derived mainly from statistical analyses of historical earthquakes and sometimes from other seismotectonic measures such as slip rates or paloseismological data. Characteristic earthquake models often rely on implicit assumptions but as often don't hold up against rigorous statistical testing (Jackson & Kagan, 2011) Earthquake loss modelling results are quite sensitive to the recurrence of small and frequent earthquakes. Engineering design ground-motion models, however, are less sensitive to such earthquakes because they are primarily based on moderate to larger levels of ground motion. On the other hand the rates of small to moderate events in the Gutenberg-Richter relationship are very sensitive to the treatment of the catalogue with regard to dependent events (after- or foreshocks) and also on the incompleteness of the historical/instrumental earthquakes. The truncated form of the Gutenberg-Richter relationship for both lower and upper bound magnitudes is usually used in computer codes for seismic hazard assessment (e.g. (McGuire & Arabasz, 1990)).

Declustering

The inclusion of foreshocks and aftershocks in the earthquake catalogue goes against the independence assumption proposed by the Poisson model and these events should thus be removed from the database. Several criteria have been suggested to distinguish foreshocks and aftershocks from main shocks (*e.g*. Gardener and Knopoff, 1974).

Typically the catalogues are declustered (e.g. Gardener and Knopoff 1974). It should be noted however that in the context of earthquake loss modelling the aftershocks that are generated over long time intervals would still contribute to the financial loss to a portfolio. Typically, losses that occur after 168 hours (irrespective of the magnitude of the shock) after a loss-causing event are considered additional loss-causing events. Insurance and reinsurance policies often cover losses above a certain threshold up to a pre-defined limit. The number of losses that can be ceded into these policies is also pre-agreed. It is very important therefore to consider aftershocks and clustered events since they may very well cause losses to these policies.

If the catalogue is declustered therefore using any of these methods, the model builder should consider how to account for these loss sources. The modelled losses would otherwise lack the loss contribution of seismological aftershocks. This differs considerably from a traditional PSHA.

Earthquake Catalogue Completeness Test

Historical earthquake catalogues are incomplete because of unequal observation periods for earthquakes of different sizes. The incomplete reporting of smaller magnitude earthquakes in earlier periods causes a bias in the earthquake catalogue. Geographical as well as temporal incompleteness usually results in inhomogeneous earthquake catalogues. In order to construct the recurrence relationship, it is therefore important to test earthquake catalogues for completeness first. For example, Stepp (1971) proposed a statistical method to correct for the incompleteness in the earthquake catalogue based on determination of a subinterval of the total time period of a catalogue, in which the mean recurrence rate is stable for a particular class of magnitude.

Magnitude conversion

An instrumental catalogue typically consists of various magnitude scales. Ultimately the magnitude scale used has to correspond to the magnitude scale (typically Mw) used by the GMPE's (see later). Magnitude conversions – particularly for local magnitudes – have to be specific to the study region and should not be imported from another target region.

Earthquake Recurrence Relationship

There is a lot of scientific debate about the appropriate magnitude-frequency behavior of single sources. There is a common agreement that magnitude-frequency behavior over large areas represents a Gutenberg-Richter behavior. The earthquake recurrence relationship is the basis of the probabilistic seismic hazard analysis. The calculated recurrence rate and the b-values for a set of earthquake magnitudes are subject to strong variations, depending on the fitting technique used. A comparison between two of the more common techniques in use, the least squares and the maximum likelihood procedures, has been made by Weichert (1980) and Bender (1983). The ability of the maximum likelihood technique to overcome the problem of magnitude intervals with zero observations is one of the advantages of this method over the least squares technique. In the least squares method, the empty magnitude interval results in duplication of points in the regression analysis. For a well defined data set, as concluded by Weichert (1980), the unpleasantness of empty magnitude intervals and large scatter in the least squares method can be partly overcome by fitting and plotting the data on a cumulative plot.

For individual seismic sources there are arguments that they follow a characteristic earthquake model or linear power-law distribution. See e.g. (Parsons & Geist, 2009).

The model developer has to make an assessment about which behavior demonstrates the stronger scientific evidence (Kagan, Jackson, & Geller, 2012). Whatever the behavior, the uncertainties in the resulting rates have to be captured – which is more challenging using a traditional earthquake model. If both models have similar evidence a logic tree capturing this epistemic uncertainty might be desirable.

The renormalized form of Gutenberg-Richter relationship after truncation for both lower and upper bound magnitudes is usually used in computer codes for seismic hazard assessment (OpenSHA, OpenQuake, EZ-Frisk).

Minimum Magnitude

Minimum magnitude is introduced into seismic hazard assessments because of its contribution to two aspects of hazard evaluation. Firstly, it is closely related to the level of completeness of data and indicates a level below which the seismicity data base is incomplete. Allowing data below this magnitude level to be incorporated into statistical analysis causes the number of large events to be overestimated in the recurrence relationship. Secondly, it is representative of a probable lower bound, below which there is no engineering interest or no associated damage to the built environment. This is also a limitation imposed by most attenuation relationships and simply implies the minimum value below which the attenuation relationships are not valid.

Maximum Magnitude

Seismic hazard studies should include some indication of the maximum size of earthquake that might occur in the region. In probabilistic seismic hazard analysis, the maximum magnitude refers to the maximum possible earthquake (MPE), which defines an upper bound to earthquake size associated with a specific seismic source. The upper limit of magnitude (M_u), follows from the notion that a limit must exist for the amount of energy that can be stored within the Earth's crust, before it ruptures. In contrast to the lower-bound magnitude, the upper-bound magnitude is dependent on the seismicity of the region. For short return periods, the choice of maximum magnitude is not a critical factor, because the predicted ground-motions for short return periods are dominated by low-to-moderate magnitude events. The maximum magnitudes are estimated from historical observations, from tectonic characteristics, or, where data are sufficient, from statistical processing of the earthquake data. The maximum magnitude used in probabilistic seismic hazard assessment indicates the severity and frequencies of ground motion at long return periods. Use of linear extrapolation of frequency-magnitude relationships to determine the maximum earthquake in a region has been questioned (*e.g.* (Schwartz & Coppersmith, 1984)). The suitability of different methods depends on the seismotectonic characteristics of the region.

	Method	**Data Required**	**Limitation**
1	Addition of increment to largest historical magnitude (usually 1/4 to 1/2 unit)	• Historical seismicity data • Assessment of the association of seismicity with seismic sources	• Usually short historical record • Uncertainties in recurrence models
2	Extrapolation of frequency-magnitude curves	• Earthquake recurrence data from historical/instrumental data • Area of seismic source being evaluated	• Reliant on recurrence data and associated database • Calibrated only to Charleston and New Madrid regions

3	Statistical (extreme-value) treatment of seismicity data	• Seismicity catalogue • Assumed model of earthquake occurrence (e.g. Poisson, exponential magnitude distribution)	• Historical record provides very poor statistical constraints on largest earthquakes
4	Strain rate / moment rate	• Strain rate cover region or seismic source • Recurrence (frequency-magnitude) relationships	• Strain rates poorly constrained in SCI • Recurrence relationships subject to short period of record
5	Maximum seismic source dimensions	• Inferred dimension from seismicity data • Other evidence for source dimensions (geologic, strain volume, etc.)	• SCI seismicity data usually define seismic sources poorly • SCI scaling relations between dimensions and magnitude are subject of current debate
6	Analogy to other SCI areas in global database	• Global database of large SCI earthquakes • Correlation of SCI earthquake sized with seismic source characteristics	• Must characterise seismic source according to characteristics in database

Table 6-1: Methods for assessing the maximum magnitude earthquake for a seismic source within a Stable Continental Interior (SCI) region (After Coppersmith and Youngs, 1989).

Overlap of seismic source models

Where seismic activity in a given region is modelled by a combination of fault source, area source and background seismicity, care must be taken that the rates that are determined from these different models do not overlap, which would cause seismicity to be overestimated or that there is a gap in the modelled area data, which would cause an *under*estimation of future seismicity.

Computational Challenge

The nature of earthquake occurrence is that lower magnitude events (that still can be damaging) are by far more frequent than large magnitude events. This in itself presents a complex computational challenge. If it was required that in a 20km grid at least one M7.0 earthquake was generated, roughly 100 M5.0 events would need to be present in the synthetic catalogue. Even for a small country like New Zealand therefore, 70,000 rupture events would be required. These events would then have to be present multiple times in order to account for the epistemic uncertainties in the hazard assessment. For example, the logic tree for the epistemic uncertainties for ground-motion prediction equations (see below) would require this number to be 5-fold larger. In addition to these hazard uncertainties, the vulnerability uncertainty has to be considered too (see later sections). The model developer therefore not only has to be astute in sampling the earthquake occurrence adequately but also has to bear the ultimate size of the stochastic catalogue in mind.

Temporal behavior

Some earthquake catalogues show whether a seismic event exhibits temporal dependence. The seismic history of single faults in particular exhibit this kind of dependent behaviour. Several factors such as accumulated strain, shearing capacity, slip rates, tectonic stress and displacement over the interface area are believed to affect the size and recurrence rate of earthquakes in a given area. This would suggest a dependency on at least two initial conditions; the size and the elapsed time of the last earthquake. If such recurrence or renewal models are applied, various statistical methods can be adopted – both in terms of probability distribution as well as in statistical fitting procedures (Parsons T., 2008). The choice of such a model however should always be based upon strict scientific evidence rather than implicit assumptions (Kagan, Jackson, & Geller, 2012)

Ground Motions

GMPEs relate the magnitude and other source and site terms of an earthquake to expected values of shaking at a given location and are based on statistical models of available data, with or without physical boundary conditions. Because GMPEs are a statistical representation it is extremely important to treat them as such: i.e. considering and quantifying not only their median predictions but also their associated uncertainties.

It should be stressed here that the largest uncertainty in seismic risk assessment is not in the magnitude recurrence behavior but actually in how the ground-motions are calculated for a given rupture. The entire topic is very closely linked the vulnerability module and cannot be treated independently from it. The setup therefore of the vulnerability module influences the modelling choice of the ground-motion prediction equations (GMPEs) and vice-versa.

The chosen ground-motion parameters, such as the PGA, PGV, and PSA for the natural vibration period Tn, are often modeled either implicitly or explicitly as a lognormal variate in developing empirical attenuation relations. A typical attenuation relation is expressed as:

$$\ln Y_{ij}(T_n) = f\left(M_j, R_{ij}, \lambda_{ij}, T_n\right) + \eta_j(T_n) + \varepsilon_{ij}(T_n)$$

where $Y_{ij}(T_n)$ is the ground-motion parameter at the *i*th site for the *j*th seismic event, $f\left(M_j, R_{ij}, \lambda_{ij}, T_n\right)$ is the attenuation relation that is a function of the earthquake magnitude M, distance R, and a set of other explanatory variables λ; $\eta_j(T_n)$ is the inter-event variability with zero mean and standard deviation $\sigma_\eta(T_n)$; and $\varepsilon_{ij}(T_n)$ is the intraevent variability with zero mean and standard deviation $\sigma_\varepsilon(T_n)$. $\eta_j(T_n)$ and $\varepsilon_{ij}(T_n)$ are often assumed to be independent and to be normally distributed.

Treatment of uncertainty

In the deterministic method the effect of maximum magnitude and the chosen attenuation relationship immediately show their impact upon the calculated hazard. In the probabilistic method however, the influence of these factors cannot be predicted and therefore the integrative nature of this method (in which a range of earthquake sizes, their distance to the site, their probability of occurrence and the probability of producing a certain level of ground-motion) has to be included in the analysis. This contrasts with the deterministic approach in which only the maximum credible earthquake at a certain distance to the site and its deterministic estimation of ground-motion are considered.

GMPEs have an epistemic uncertainty: lack of knowledge. Even using the same dataset different authors will come up with different statistical models. Typically this source of uncertainty is modelled via a logic tree approach.

GMPEs also have an aleatory uncertainty: randomness of nature. ((Bommer & Abrahamson, 2006), (Atik, et al., 2010)). That is to say, given all measurable quantities of an earthquake the intensities at a given location cannot be precisely predicted. The standard deviations of the ground-motion residuals are typically split into a component of intra-event uncertainty and inter-event uncertainty. The inter-event uncertainty is a residual per event, the intra-event uncertainty a residual per event and location.

It is key to understand that the inter-event component is a correlated aspect that affects all sites of interest at the same time. Although it is the smaller of all the total aleatory uncertainties it can have a much larger influence on the total loss since all sites in a portfolio would be affected by higher than average ground-motions at the same time (e.g. (Bazzurro & Luco, 2005)).

Note also the dependency to the vulnerability module at this point. Claims data will always be deficient in that the intensity at the site of loss will only approximately be known. In high-resolution claims data there will be however an overlap of the variability (and inferred uncertainty) of damage degree versus the variability in ground motions – hence resulting in a double-counting of uncertainty. Great care must be taken to ensure that the ground-motion variability is not double-counted (see e.g. (Crowley, Bommer, & Stafford, 2008)).

Spatial correlation and period correlation

- Note the intra-event uncertainty is spatially correlated (e.g. (Loth & Baker, 2013))
- Note that there is a correlation across multiple natural vibration periods T_n(e.g. (Goda & Hong , 2008))

Regional applicability

Regional and global GMPEs are typically available when considering the seismic hazard of a region and selection must be done with care. The initial and apparently most applicable one might just fit the dataset as observed thus far best but might neglect knowledge gained from other similar tectonic regions. Statistical overfitting must be borne in mind if considering regional GMPEs (Bommer, et al., 2010).

Amplification

Local site conditions influence the ground-shaking. The amplification of seismic waves is often dealt with by means of proxy variables – vs30 (average shear wave velocity in the upper 30 meters) being the most common. Vs30 measurements are typically only available at few specific sites and seismic stations. A simple interpolation between those stations is not adequate. Hence proxy variables such as topographic slope might be used (see e.g. (Wald & Allen, 2007)) to evaluate the amplification potential at many potential sites of interest.

Amplification is highly dependent on the ground motion intensity measure used and can differ depending on the input ground motions level. For low vs30 soils and high-frequency ground motions the rock-site ground motions might be de-amplified. (e.g. (Field, Johnson, Beresnev, & Yuehua, 1997))

Hazard Benchmarking

The outcome of the hazard model should be compared to other studies. Hazard maps and hazard curves are typically available for a comparison. Local and regional differences should be explained. One can consider comparing hazard curves to actual reported and/or measured ground motions, although this is a challenging task (e.g. (Stirling & Gerstenberger, 2010))

Vulnerability

Vulnerability links the modelled intensity measure to a financial loss and is dependent on the risk characteristics of interest. For a property portfolio of interest to an insurer these are typically construction type, occupancy, age of applied seismic code, building practices and construction quality and building height. Depending upon the application other parameters might also be relevant.

In practice it will be impossible to obtain all relevant risk characteristics for each and every risk in the portfolio. Careful consideration therefore has to be made as to how the model should react in case not all risk parameters are provided.

The relationship between intensity measure and a damage degree is non-linear. Damage is usually expressed as a damage degree, the ratio of absolute financial loss to the replacement value of the property (TIV, total insured value). The reason for expressing the damage as a damage degree is that once a robust relationship is developed financial conditions such as deductibles and limits can easily be applied. Most typically deductibles are expressed as a fraction of TIV s capturing the damage states as fractions of TIV as well is very practical. This further ensures that the vulnerability functions developed for one country or region can be applied elsewhere.

Vulnerability not only links the intensity measure of interest to expected damage degree but also assesses the uncertainty of the distribution around the expected value. The key benefit of a probabilistic risk framework is giving quantitative information on the uncertainty. Vulnerability is the key uncertainty in seismic risk assessment.

What Intensity measure?

What is the object of loss being considered? What correlates best with actual losses? Typically an insured portfolio will consist of a mix of various risks. Each risk should be assessed for whether a short-period natural vibration period might be more appropriate (low-rise buildings) or a long-period one (tall buildings).

Importance of claims history

The most robust way to build vulnerability functions is by means of claims history. As vulnerability functions describe financial loss, they are influenced by parameters outside the simple structural performance of a building. How expensive rebuilding material is, the workforce used to rebuild and even the legislative system in a country may all influence the expense of reconstruction. It is clear that it is very important to have actual financial claims data.

Engineering-based approaches, such as using fragility curves, yield physical damage, not monetary loss to assets. In the absence of any claims data engineering-based approaches can be useful, although great care must be taken in calculating how the structural damage translates to financial loss. Engineering-based studies are however very useful in determining the relativities between two risk characteristics. For example, there might be abundant claims information about single-storey wood-frame buildings but little available on reinforced concrete tilt-up structures. In this case engineering considerations become very important in differentiating between these two risk categories.

When considering claims history it is also important to gather the portfolio values at the time of loss. Even in an earthquake most building are typically not damaged and there is no resulting financial loss. Capturing how many buildings are *not* damaged is an important aspect of getting the impact of financial conditions correct. In practice the replacement values of individual risks with losses are not reported. The claims information therefore must be linked to the portfolio information in order to estimate a damage degree. If the replacement value is known it must be assessed whether it is true replacement value and not, for example, net of insurance conditions. Sources of loss in the claims data that are not modelled explicitly should be removed from the claims data. However one has to assess and document clearly It is important to identify which loss sources have been removed and are therefore not included in the final model.

Other aspects

- Note that the setup of the vulnerability module also has a dependency on how the GMPEs are implemented. Aleatory variability could from a conceptual perspective also be incorporated into the vulnerability module, so that the distribution in damage degree reflects the joint probability distribution of ground motion and damage degrees.
- Claims experience will never be complete or incorporate all risk characteristics that are relevant for a portfolio of assets. In order to enrich claims experience engineering studies are still very important. The knowledge of relative risk performance should be

incorporated in an earthquake loss model. If the relative vulnerability of risks is not incorporated this might lead the user to anti-select risk in relation to its peers.
- Current portfolio information does not incorporate all building characteristics. In the absence of these, meaningful average values have to be developed. The earthquake risk model has to deliver a meaningful average building stock that is representative for classes of risk in the market.
- Design practices and design codes vary regionally – not only at a national level. Such regional differences have to be incorporated.
- The shape of the vulnerability curve and the shape of the probability distribution of it will greatly influence the deducable credit of the loss model. The condition credit or deductible credit describes how much probability is allocated to a given loss level. This has a profound impact on how financial conditions are treated in the loss model. For example, the more probability is allocated to losses with high damage degrees, the lower the impact of typical financial structures, such as deductibles, are.

Further loss components

Post Loss amplification

After major events it can be seen that there is a loss component that does not stem from individual assets but rather from the fact that many assets are affected at the same time. Often prices for reconstruction material become more expensive and hence inflate the financial loss: economic demand surge. Claims either cannot be settled or can only with delay, or the number of claims exceeds the claims handling capabilities of the insurers: claims inflation.

Non-modelled hazards

It should be noted that even when most physical impacts from an earthquake, such as the shaking losses, losses from fire-following, tsunami, landslide and liquefaction are considered there are other loss sources which cause financial loss but which are not modelled. Such loss sources can include, but are not limited to, debris removal costs, looting, pipe-burst etc.

Loss Benchmarking

Even the most sophisticated earthquake risk model should be consistent to loss experience. In other words, how do modelled event losses compare to actual event losses on a portfolio level? How often is a specific loss level exceeded in observed history and what frequency of such loss levels is returned by the model?

BIBLIOGRAPHY

Ambraseys, N.N. and Melville C.P. [1982]. *A history of Persian earthquakes*, Cambridge University Press, Cambridge, 219 pp.

Ambraseys, N.N. and Melville C.P. [1996] *The catalogue of Persian Earthquakes*, Imperial College of London

Atik, L. A., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N. (2010). The Variability of ground-motion prediction models and its components. *Seim. Res. Let.* , *81*(5), 794-801.

Bakun , W. H., Johnston , A. C., & Hopper , M. G. (2003). Estimating Locations and Magntiudes of Earthquakes in Easter North America from Modified Mercalli Intensities. *Bulletin of the Seismological Society of America, 93*(1), 190-202.

Bazzurro, P., & Luco, N. (2005). Accounting for uncertainty and correlation in earthquake loss estimation. Rome: Proceedings of the 9th International Conference on Structural Safety and Reliability (ICOSSAR).

Bommer, J. J., & Abrahamson, N. A. (2006). Why do modern probabilisitc seismic-hazard analyses often lead to increased hazard estimates? *Bull. Seism. Soc. of Am. , 96*(6), 1967-1977.

Bommer, J. J., & Scherbaum, F. (2008). The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis. *Earthquake Spectra, 24*(4), 997–1009.

Bommer, J. J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H., & Fäh, D. (2010). On the selection of ground-motion prediction equations for seismic Hazard analysis. *Seismological Research Letters, 81*(5).

Crowley, H., Bommer, J. J., & Stafford, P. J. (2008). Recent Developments in the Treatment of Ground-Motion Variability in Earthquake Loss Models. *Journal of Earthquake Engineering*, 71-80.

Field, E. H., Biasi, G. P., Bird, P., & et al. (2013). Uniform California earthquake rupture forecast, version 3 (UCERF3)—The time-independent model: U.S. Geological Survey Open-File Report 2013. *California Geological Survey Special Report 228*. Retrieved from http://pubs.usgs.gov/of/2013/1165/

Field, E. H., Johnson, P. A., Beresnev, I. A., & Yuehua, Z. (1997). Nonlinear ground-motion amplification by sediments during the 1994 Northridge earthquake. *Nature, 390*, 599-602.

Frankel, A. (1996). National Seismic Hazard Maps: Documentation June 1996. *USGS Open-File Report 96-532*.

Goda, K., & Hong , H. P. (2008). Spatial Correlation of Peak Ground Motions and Response Spectra. *Bulletin of the Seismological Society of America, 98*(1), 354-365.

Hayes, G. P., Wald, D. J., & Johnson, R. L. (2012). Slab1.0: A three-dimensional model of global subduction zone geometries. *J. Geophys. Res.* doi:doi:10.1029/2011JB008524

Hiemer, S., Woessner, J., Basili, R., Danciu, L., Giardini, D., & Wiemer, S. (2014). A smoothed stochastic earthquake rate model considering seismicity and fault moment release for Europe. *Geophys. J. Int. , 198*, 1159–1172.

Jackson, D. D., & Kagan, Y. Y. (2011). Characteristic earhtquakes and seismic gaps. *Encyclopedia of Solid earth geophysics*, 37-40.

Kagan, Y. Y., Jackson, D. D., & Geller, R. J. (2012). Characteristic Earthquake Model 1884-2011, R.I.P. *Seismological Resarch Letters, 83*(6), 951-953.

Loth, C., & Baker, J. W. (2013). A spatial cross-correlation model of ground motion spectral accelerations at multiple periods. *Earthquake Engineering & Structural Dynamics, 42*(3), 397–417.

McGuire, R. K., & Arabasz, W. J. (1990). An introduction to seismic hazard analysis. *In: S. H. Ward Geotechnical and Environmental Geophysics,*, 333-353.

Musson, R. M. (2000). The use of Monte Carlo simulations for seismic hazard assessment in the U.K. *Annali di Geofisica, 43*(1).

Parsons, T. (2008). Monte Carlo method for determining earthquake recurrence parameters from short paleoseismic catalogs: Example calculations for California. *U.S. Geological Survey Open-File Report 2007-1437C and California Geological Survey Special Report 203C*.

Parsons, T., & Geist, E. L. (2009). Is There a Basis for Preferring Characteristic Earthquakes over a Gutenberg–Richter Distribution in Probabilistic Earthquake Forecasting? *Bulletin of the Seismological Society of America*, 2012–2019.

Schwartz, D. P., & Coppersmith, K. J. (1984). Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas Fault Zones. *J. Geophys. Res. , 89*(B7), 5681–5698 .

Stirling, M., & Gerstenberger, M. (2010). Ground Motion–Based Testing of Seismic Hazard Models in New Zealand. *Bulletin of the Seismological Society of America, 100*(4), 1407–1414.

Storchak, D. A., Giacomo, D., Bondàr, I., Engdahl, E. R., & et al. (2013). Public Release of the ISC-GEM Global Instrumental Earthquake Catalogue (1900-2009). *Seismological Research Letters, 84*(5).

Strasser, F. O., Arango, M. C., & Bommer, J. J. (2010). Scaling of the Source Dimensions of Interface and Intraslab Subduction-zone Earthquakes with Moment Magnitude. *Seim. Res. Lett., 81*(6), 941-950. doi:10.1785/gssrl.81.6.941

Wald, D. J., & Allen, T. I. (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bull., Seism. Soc. Am, 97*(5).

Wells, D. L., & Coppersmith, K. J. (1994). New Empirical Relationships among Magntiude, Rupture Lenght, Rupture Width, Rupture Area, and Surface Displacement. *Bulletin of the Seismological Society of America, 84*(4), 974-1002.

Zolfaghari M.R. (2015) "Development of a synthetically generated earthquake catalogue towards assessment of probabilistic seismic hazard for Tehran, Journal of Natural Hazard.