Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ETH RISK CENTER

LECTURE NOTES

Fundamentals of Probabilistic Seismic Hazard Analysis (PSHA)
Chapter 2: Lecture 4

Spring 2018

Written by simona esposito
marco broccardo
IBK, ETHZ

"Every earthquake contains a surprise"
Clarence Allen, 1952

Abstract: These lecture notes present the main aspects of Probabilistic Seismic Hazard Analysis (PSHA). PSHA is the basis for many earthquake design codes practices, and for seismic risk analysis of civil systems. The study of seismic hazard requires an understanding of the various processes by which earthquakes occur and their effects on ground motion. Therefore, in the first part of this note, we provide a brief introduction to the terminology used to describe the complex phenomena of earthquakes. Then, each of the steps needed to perform a site specific PSHA is presented and discussed.

1 Introduction

Seismic hazard analysis is one of the main components for the evaluation of the seismic risk at an earthquake prone site. In particular, the seismic hazard can be defined as the exceedance (or occurrence) probability for a given ground motion intensity measure threshold, site and time interval. Seismic risk is the exceedance (or occurrence) probability for a given loss threshold, site and time interval.

Reliable estimates of a seismic hazard are paramounts to minimize loss of life, property damage and social and economic disruption. Further, seismic hazard analysis is the base-ground for seismic-structural design, financial-insurance policies, emergency-response policies, and post-earthquake recovery.

One of the main use of hazard analysis is within the context of seismic design of buildings and infrastructures. These results are usually incorporated into building codes for residential buildings; while for critical facilities, hazard analysis are performed ad hoc.

The approaches to seismic hazard assessment can be grouped into two broad categories: deterministic and probabilistic.

A Deterministic Seismic Hazard Analysis (DSHA) consists in the analysis of a particular seismic scenario. Generally, the output of the analysis is the intensity level of the selected hazard scenario. In particular, the scenario consists of the postulated occurrence of an earthquake of a specific size occurring at a given location. A typical DSHA (shown in Figure 1) can be split in four stages:

i Identification and characterization of all earthquake sources capable of producing significant ground motion at the site. Source characterization includes definition of the earthquake source's geometry (source zone) and earthquake potential.

ii Selection of a source to site distance parameter for each source zone. Usually the shortest distance between the source zone and the site of interest is selected. The distance can be expressed in different ways (e.g. epicentral or hypocentral) depending on the measure of distance of the predictive equation used in the following step.

iii Selection of the controlling earthquake (i.e. the earthquake that is expected to produce the strongest level of shaking) generally expressed in term of ground motion parameters at the site. The selection is made comparing the levels of shaking produced by earthquakes (identified in step i) assumed to occur at the distances identified in steps ii. The controlling earthquake is described in terms of its size (i.e. magnitude) and distance from the site.

iv The hazard at the site is formally defined, usually in terms of the ground motions produced at the site by the controlling earthquakes. Its characteristics are usually described by one or more ground motion parameters.

The DSHA provides a straightforward framework for the evaluation of the worst-case ground motions. This is a useful procedure when applied to critical facilities (e.g. nuclear power plant), where failure could have catastrophic effects. However, it does not provide any information regarding the likelihood of occurrence of the controlling earthquakes, the level of shaking that maybe expected during the lifetime of a particular structure, and the uncertainty involved in all the steps of the procedure. Finally, it is important to recognize that at the step i (but not only), DSHA involves critical subjective decisions.

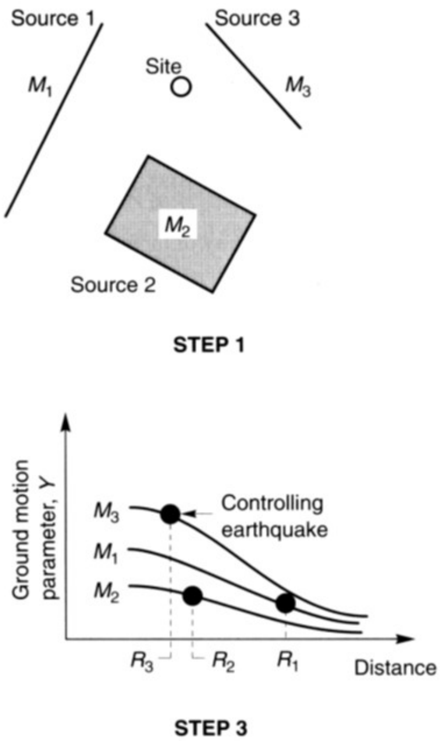

Figure 1: Four steps of a DSHA analysis (Kramer, 1996).

Probabilistic seismic hazard analysis (PSHA) provides an uncertainty quantification framework to address the hazard analysis. PSHA is an essential tool for current earthquake engineering practices and performance-based earthquake engineering (PBEE) design methodology. The goal of a PSHA is to evaluate the exceedance (or occurrence) probability of a given ground motion intensity measure threshold at a considered site and in a considered time interval. These intensity measures are then related to facilities damages, economic and social losses.

The numerical/analytical approach to PSHA was first formalized by Cornell (1968). The most comprehensive treatment to date is the SSHAC (1997) report, which covers many important procedural issues that will not be discussed here. The SSHAC report represents the best source of additional information for anyone conducting a PSHA.

The PSHA cans also described in four steps similarly to the steps of a DSHA procedure as illustrated in Figure 2:

i The first step is identical to DSHA, except that the probability distribution of potential earthquake locations within the source must also be characterized. These distributions are then combined with the source geometry to obtain the corresponding probability of source to site distance.

ii The temporal distribution of earthquake recurrence and size distribution (magnitude) must be characterized for each source zone. A recurrence relationship specify the average rate at which an earthquake of some size will be exceeded.

iii The ground motion produced at the site by earthquakes of any possible size occurring at any possible point in each source zone must be determined with use of predictive equations.

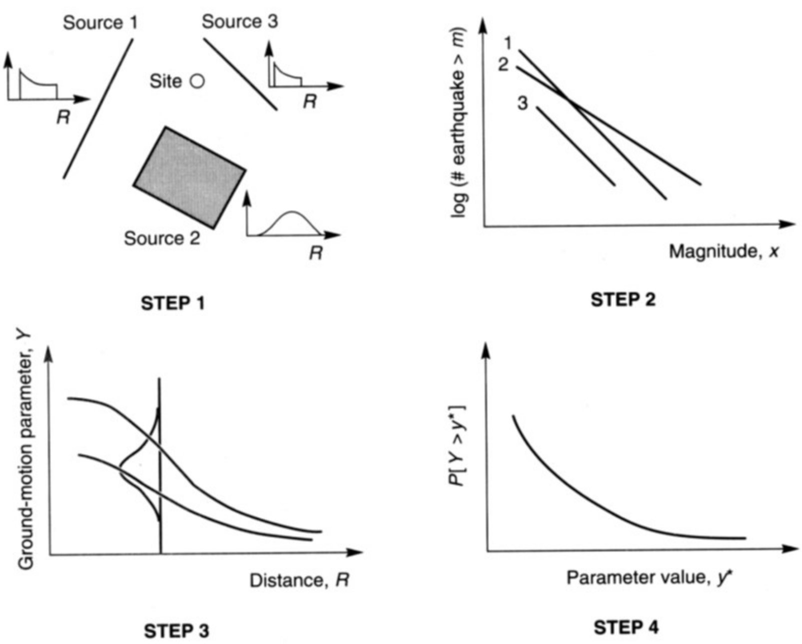

Figure 2: Four steps of a PSHA analysis (Kramer, 1996).

iv In the last step the uncertainties in earthquake location, earthquake size, and ground motion are combined to obtain the probability that the ground motion parameter will be exceeded during a particular time period.

The selection of the most appropriate approach depends on several factors including the seismic environment, the structures and the aim of the project. However, "the more quantitative the decision to be made, the more appropriate is probabilistic hazard and risk assessment" (Mc Guire, 2004).

PART I
Seismology in a Nutshell

2 Earthquakes

Earthquakes are the result of a sudden release of energy in the Earth's crust that creates elastic waves: the Earth reacts as an elastic solid and seismic waves are propagated to all parts of the Earth following paths through the body of the Earth itself and around its surface.

Earthquakes may be classified as natural or anthropogenic according to the nature of the source. Anthropogenic earthquakes are induced by human activity, such as technological operations involving extraction and/or injection of fluids in underground rocks. Natural earthquakes are caused by natural processes in the Earth and their nature can be volcanic or tectonic.

Earthquakes produced by stress changes in solid rock due to the injection or withdrawal of magma are called volcano earthquakes. These earthquakes can cause land to subside and can produce large ground cracks. These earthquakes can occur as rock is moving to fill in spaces where magma is no longer present. Volcano earthquakes don't indicate that the volcano will be erupting but can occur at any time. In this type of earthquakes the direct cause is an induced effect of the geo-dynamic process. In tectonic earthquakes the direct cause is the geodynamic movement itself.

Earthquakes occurring at boundaries of tectonic plates are called interpolate earthquakes, while the less frequent events that occur along faults in the normally stable interior of the lithospheric plates are called intraplate earthquakes. Intraplate earthquakes often occur at the location of ancient failed rifts, because such old structures may present a weakness in the crust where it can easily slip to accommodate regional tectonic strain. These earthquakes are not well understood, and the hazard associated may be difficult to quantify.

Most tectonic earthquakes are causally related to compressional or tensional stresses built up at the margins of the lithospheric plates. These occur when rocks in the earth's crust break due to geological forces created by movement of plates. Boundaries have different names depending on how the two plates are moving in relationship to each other (Figure 3):

i Crashing: convergent (subduction zone) boundaries.

ii Pulling apart: divergent (or spreading ridge) boundaries.

iii Sideswiping: transform fault boundaries.

Plate movement causes the buildup of tremendous quantities of energy in the rock. When the rock's rupture strength is exceeded, the stored energy is suddenly released producing vibrations that travel through the rock, leading to earthquakes.

2.1 Faults

In some regions, plate boundaries are easy to identify, while in others they may consist in smaller plates or microplates trapped between larger plates. Locally the movement between two portions of the crust will occur in new or preexisting offsets in the geological structure of the crust known as *fault*.

Faults are planar rock fractures which show evidence of relative movement. Earthquakes are caused by energy release during rapid slippage along faults. The largest examples are at tectonic plate boundaries, but many faults occur far from active plate boundaries. Since faults usually do not consist of a single, clean fracture, the term fault zone is used when referring to the zone

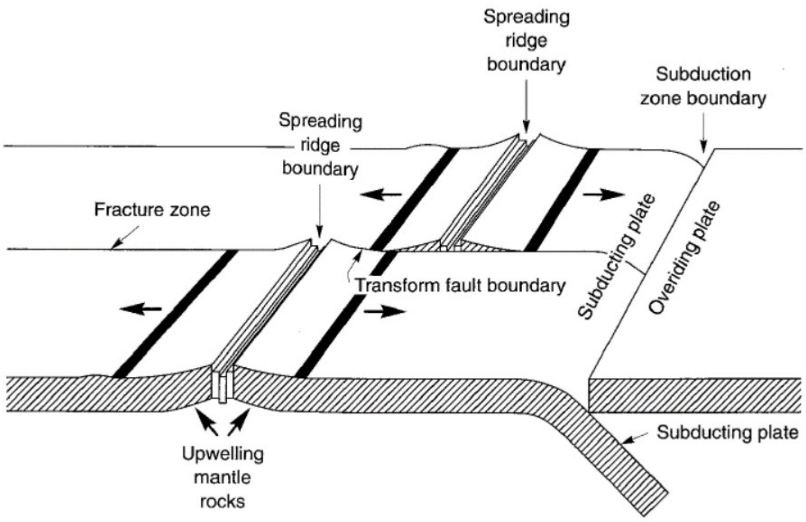

Figure 3: Plate boundaries (Kramer, 1996).

of complex deformation associated with the fault plane. The fault plane is then defined as the surface where the movement taken place within the fault.

The two sides of a fault are called the *hanging wall* and *footwall.* The hanging wall occurs above the fault. The footwall occurs below the fault.

The geometry of the fault plane is described by its strike and dip (Figure 4). The strike is the horizontal line produced by the intersection of the fault plane and a horizontal plane. The strike azimuth describes the orientation of the fault with respect to the north. The dip angle is the angle between the fault plane and the horizontal plane measured perpendicular to the strike.

The *sense of slip*, defined by the relative movements of geological features present on either side of the fault plane, defines the type of fault. Three groups are identified (5):

i *Dip-slip fault*: where the main sense of movement (or slip) on the fault plane is vertical.

ii *Strike-slip (or transform) fault*: where the main sense of slip is horizontal the fault is known as a fault.

iii *Oblique-slip fault*: has significant components of both strike and dip slip.

Dip-slip faults include both normal and reverse. A *normal fault* occurs when the crust is in extension. The hanging wall moves downwards relative to the footwall. A *reverse fault* is the opposite of a normal fault: the hanging wall moves up relative to the footwall. Reverse faults are indicative of shortening of the crust. The dip angle of a reverse fault is relatively steep, greater than 45°. A *thrust fault* has the same sense of motion as a reverse fault, but with the dip of the fault plane at less than 45°.

In the strike-slip faults, the fault surface is usually near vertical and the footwall moves either left or right or laterally with very small vertical motion. The San Andreas fault is a remarkable example of a strike-slip fault. It marks the boundary between the North American and Pacific Plates in California.

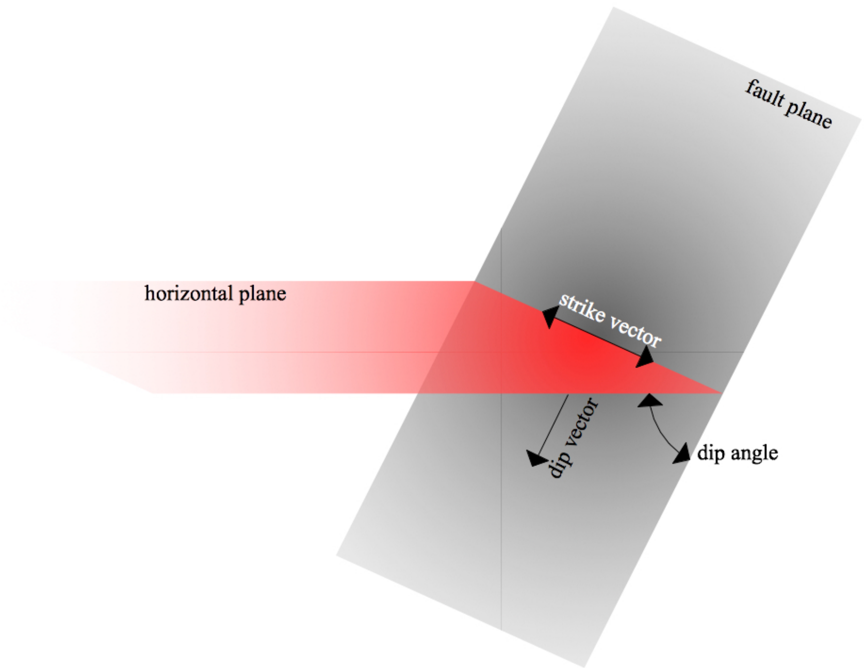

Figure 4: Fault plane orientation. Modified from Kramer, 1996.

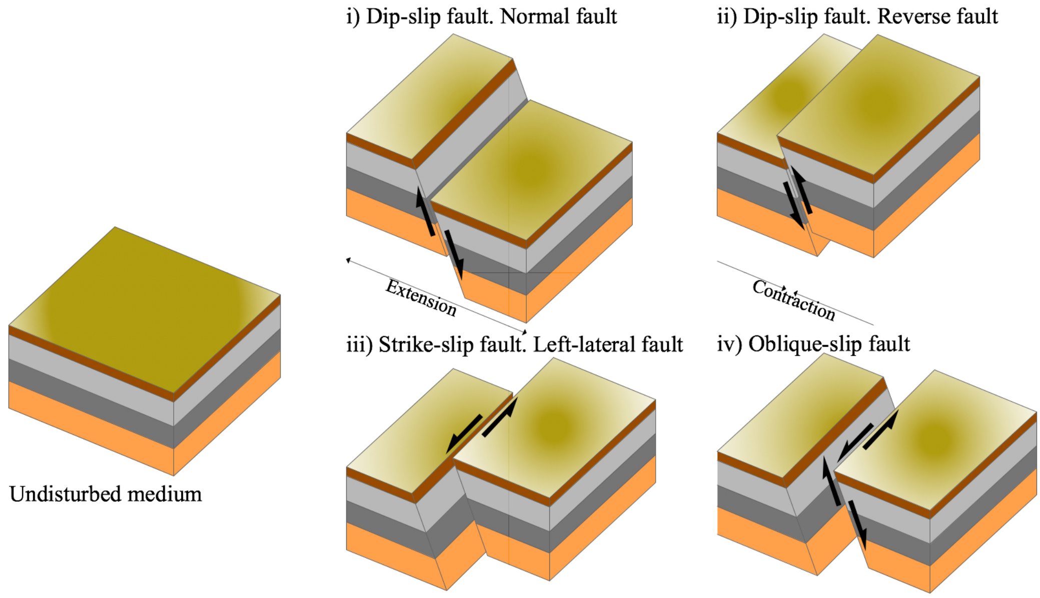

Figure 5: Type of faults

2.2 Seismic waves propagation

When an earthquake occur, seismic waves radiate away from the source to the ground surface and travel rapidly though the earth's crust. Reaching the surface, seismic waves produce shaking which strength and duration depends on the size and location of the earthquake and on the characteristics of the site.

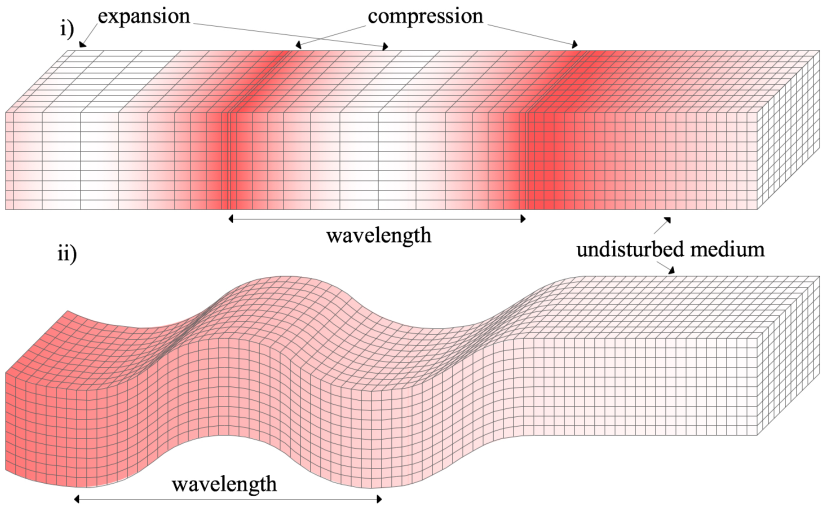

Figure 6: Deformations produced by body waves: i) P- waves ii) S-V waves.

The ground shaking felt at a given location will be made up of a combination of *body waves* and *surface waves*.

Body waves propagate through earth and they are generating by seismic faulting while surface waves travel along the ground surface and in most of cases are generated by the reflection and refraction of body waves. Body waves (Figure 6) include compressional waves (P-waves) where the ground moves parallel to the direction of propagation, and shear waves (S-waves) where the ground moves perpendicular to the direction of propagation; moreover depending on the direction of particle movement, S-waves can be divided in SV (vertical plane movement) and SH (horizontal plane movement).

Surface waves (*Rayleigh* and *Love waves*, Figure 7), instead, are more complex: for Rayleigh waves the particle motion traces a retrograde ellipse in a vertical plane with the horizontal component of motion being parallel to the direction of propagation; for the Love waves, the particle motion is along a horizontal line perpendicular to the direction of propagation.

The amplitude of ground motion reduces with distance from the source of seismic energy release. This is due to a combination of geometric attenuation, which accounts for the spread of the wave front as it moves away from the source, and anelastic attenuation, which is caused by material damping. In the immediate locality of the fault rupture, body waves will dominate the motion while ground motion at large distances to the source is generally dominated by surface waves because of the geometric attenuation is different for the two types of waves: assuming that the earthquake rupture zone may be represented as a point source and R is the distance from the rupture zone, the amplitude of body waves decreases in proportion to $1/R$, while the amplitude of surface waves decreases in proportion to $1/\sqrt{R}$.

2.3 Site effect

Although seismic waves travel through the rock for the majority of their trip from the source to the ground surface, the final part of the trip is through soil, which characteristics may influence the nature of shaking at the ground surface. The soil tends to act as a “filter” to seismic waves attenuating or amplifying the motion. Since soil conditions over short distances, levels of ground shakings may vary dramatically also within a small area. Generally site effects represent local

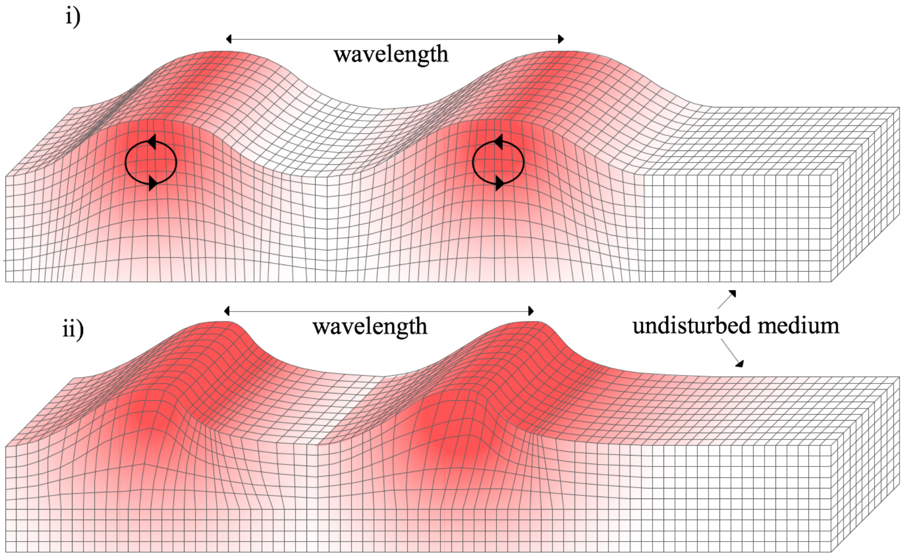

Figure 7: Deformations produced by surface waves: a) Rayleigh waves b) Love waves.

ground response effects, basin effects and the influence of surface topography on ground motion.

Local ground response refers to the influence of relatively shallow geologic materials on (nearly) vertically propagating body waves.

The term basin effects refers to the influence of two- or three-dimensional sedimentary basin structures on ground motions, including critical body wave reflections and surface wave generation at basin edges.

Site effects due to surface topography (i.e., topographic effects) can amplify the ground shaking that would otherwise be expected on level ground along ridges or near the tops of slopes.

2.4 Location of earthquakes

Earthquakes result from rupture of the rock along a fault. The point at which rupture begins and first seismic waves originate is called the hypocenter (or focus) of the earthquake (Figure 8). From this point, located at some focal depth (or hypocentral depth) the earthquake spreads across the fault at velocities of 2 to 3 km/s. The point on the ground surface above the focus is called epicenter.

The distance between a site and the hypocenter is called hypocentral distance, while the distance on the ground surface between the site and the epicenter is known as the epicentral distance.

Earthquakes can be divided into three categories according to the hypocenter depth:

i *Shallow*: maximum depth of 60 km.

ii *Intermediate*: with a depth varying from 60 to 300 km.

iii *Deep*: with a depth varying from 300 to 650 km.

The location of an earthquake is usually specified in terms of the location of its epicenter. Preliminary location of epicenter is based on the differential wave-arrival times measurements

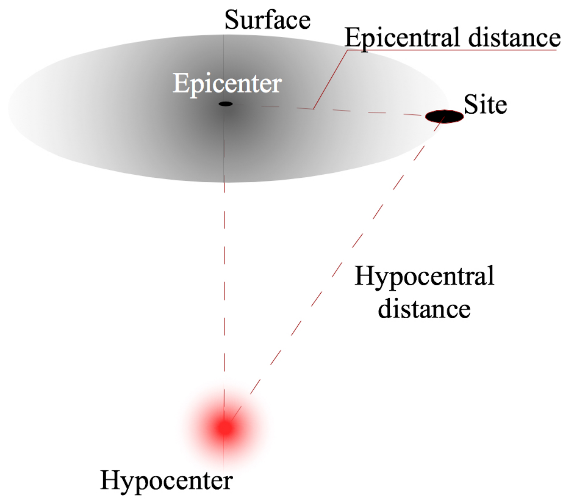

Figure 8: Earthquake location.

at a set of at least three seismographs. In particular it is based on the different arrival times of P- and S-waves. Since P-waves are faster than S-waves, they will arrive first at a given seismograph. The distance between the seismograph and the epicenter of the earthquake (d) will depend on the difference in arrival time of P-and S-waves (Δt_{p-s}) and the difference between the P- and S-wave velocities (λ_p and λ_s) according to:

$$d = \frac{\Delta t_{p-s}}{1/\lambda_s - 1/\lambda_p} \tag{1}$$

At any single seismograph it is possible to determine only the epicentral distance but not the direction of the epicenter. Plotting a circle of radius equal to the epicentral distance for a set of three seismograph, the location of the epicenter will corresponds to the intersection of the three circles.

2.5 Size of earthquakes and magnitude scales

The oldest measure of the size of an earthquake is represented by the earthquake intensity. The intensity is a qualitative description of the effects of the earthquake at a particular location as evidenced by observed damage and human reactions. The increasing diffusion of modern instrumentations made necessary an absolute and instrumental scale of the earthquake size called earthquake magnitude. Most of the scales are based on measured ground motion parameters.

The fundamental step was done by Charles F. Richter in 1935 who used a Wood-Andersen torsional seismometer to define a magnitude scale for shallow, local earthquakes in California.

The *local magnitude* or *Richter magnitude* M_L is defined as the difference between the logarithm (base 10) of the maximum trace amplitude ($\log A$) (in millimeters) recorded at a known distance and the value of the "Zero Event" ($\log A_0$) associated to the same distance:

$$M_L = \log A - \log A_0 \tag{2}$$

The "Zero Event" (and the corresponding curve) is the event producing a 0.001 mm peak amplitude if recorded by Wood-Anderson instrument 100 km far from the epicenter of the event. Usually for each seismic station, two signal are available recorded in two orthogonal horizontal direction (EW and NS): estimated local magnitude is the mean of values computed in each direction. Similarly, if more stations recorded the same event, local magnitude is the mean of values computed by each station.

Example I Evaluate the local magnitude of an event with a recorded peak amplitude of A=45 mm, located 100 km from the epicenter.

$$M_L = \log 45 - \log 0.001 = 1.66 - (-3.0) = 4.66$$

The local magnitude is the best known magnitude scale but it is not always the most appropriate scale for the description of the earthquake size. In order to extend the magnitude scale to more distant and deeper events, other magnitude scale based on the amplitude of a particular wave have been introduced. In particular two alternative magnitude scales M_S and m_b which refer to surface wave and volume (body) waves respectively, were introduced.

The *surface wave magnitude* (M_S) is based on the Rayleigh waves with a period of about 20 seconds which are not really influenced by the crust characteristics. It is obtained from:

$$M_S = \log A - 1.66 \log D + 2.0 \tag{3}$$

where A is the maximum ground displacement in micrometers and D is the epicentral distance of the seismometer measured in degrees. This magnitude is commonly used to describe the size of shallow (< 70 km focal depth) distant (Distance > 1000 km) moderate to large earthquakes.

For deeper events, volume waves have to be used because superficial waves are often too small to permit reliable evaluation of the surface wave magnitude. The magnitude scale with refers to volume P-waves (compressional and longitudinal in nature) which are not strongly influenced by the focal depth is the body wave magnitude, expressed as:

$$m_b = \log\left(\frac{A}{T}\right) + 0.01D + 5.9 \tag{4}$$

where A in this case is the P-wave amplitude in micrometers and T is the period of the P-wave.

Previous described scales are empirical quantities based on measured ground motion parameters. However ground-shaking characteristics do not necessarily increase at the same rate of the amount of energy released during an earthquake. For large earthquakes, the energy recorded at one location does not continue to increase as the earthquake rupture area increases. This phenomenon is referred to as *saturation*. M_L and m_b saturate at magnitudes of 6 to 7 and M_S saturates at around 8. The only magnitude scale that is not subjected to saturation is the moment magnitude.

The *moment magnitude* (M_W) is a measure of earthquake size that can be related to physical parameters of an earthquake and it does not saturate because of seismograph limitations. It is based on the seismic moment which is a direct measure of factors that produce rupture along the fault. It is given by:

$$M_W = \frac{2}{3}\log(m_{min}) - 10.7 \tag{5}$$

where m_{min} is the seismic moment in dyne-cm, which is a function of soil shear modulus (μ), rupture area (A) and average displacement ($\overline{\Delta u}$) on A,

$$m_{min} = \mu A \overline{\Delta u} \tag{6}$$

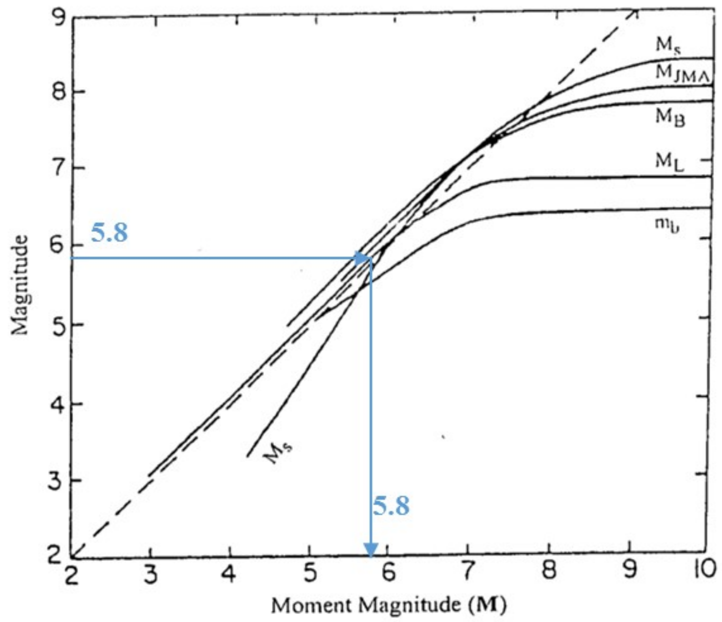

Figure 9: Relation between Moment Magnitude (M_W) and other Magnitude scales, adapted from McGuire (1994). M_{JMA} is the magnitude reported by the Japan Meteorology Agency. Note that $M_W = M_L$ for $M_W < 6.2$ and $M_W = M_S$ for $6.2 < M_W < 8.0$.

Fault type	**Relationship**	$\sigma_{\log L,A,D}$
Strike slip	$\log L = 0.74M_W - 3.55$	0.23
Reverse	$\log L = 0.63M_W - 2.86$	0.20
Normal	$\log L = 0.50M_W - 2.01$	0.21
All	$\log L = 0.69M_W - 3.22$	0.22
Strike slip	$\log A = 0.90M_W - 3.42$	0.22
Reverse	$\log A = 0.98M_W - 3.99$	0.26
Normal	$\log A = 0.82M_W - 2.87$	0.22
All	$\log A = 0.91M_W - 3.49$	0.24
Strike slip	$\log D = 1.03M_W - 7.03$	0.34
Reverse	$\log D = 0.29M_W - 1.84$	0.42
Normal	$\log D = 0.89M_W - 5.90$	0.38
All	$\log D = 0.82M_W - 5.46$	0.42

Table 1: Empirical Relationships between, surface rupture length (L), rupture area (A), maximum surface displacement (D).

Figure 9 summarizes and compares some of these magnitude scales and the saturation level of each scale.

The previous equation shows that seismic intensity is somehow correlated to the rupture dimensions. Analyzing historical earthquake data, Wells and Coppersmith (1994) provided empirical relationship between moment magnitude and all the geometrical parameters of the rupture (Table 1).

Example II Compute the probability that a moment magnitude 7.0 earthquake on the San Andreas Fault would cause a rupture area larger than 600 [km^2].
Solution: San Andrea is a strike-slip movement. From Table 1, the mean rupture area is:

$$\log A = 0.9M_W - 3.42 = 6.3 - 3.42 = 2.88$$

from which

$$A = 10^{2.88} = 758.6 \text{ [km}^2\text{]}$$

The standard normal variate for a 400 [km^2] area is

$$\begin{aligned} z &= \frac{\log 600 - \log 758.6}{0.22} = -0.46 \\ P(A > 600) &= 1 - P(A \leq 600) = 1 - \Phi(z) = 1 - 0.32 = 0.68 \end{aligned}$$

2.6 Ground motion intensity measures

The magnitude measures the intensity of a seismic event and it is a source parameter because it depends only on the characteristics of the rupture generating the event. On the contrary, it is also important to define some specific parameters able to identify the ground motion intensity in terms of effects in specific sites. Instrumental recordings of the strong ground movements during earthquakes are the clearest and most comprehensive definition of the actions against which structures and lifelines must be designed. Accelerographs record the acceleration of the ground as a function of time, while seismographs record the displacement or velocity of the ground. Double integration of the accelerogram provides the velocity and displacement time-histories to be recovered as well. Ground motion parameters (also called Intensity Measures, IMs) describe the most important characteristics of strong ground motion in compact and quantitative form. Many parameters have been proposed to characterize:

i Amplitude (i.e. how large is the shaking?).

ii The frequency content (what frequencies are particularly prevalent in the ground motion?).

iii The duration (i.e. how long does the strong shaking last?) of strong ground motions.

Some intensity measures describe only one of these characteristics, while others may reflect two or three.

The most commonly used measure of the amplitude of a particular ground motion is the horizontal *peak ground acceleration* (PGA). The PGA for a given component of motion is the largest (absolute) value of horizontal acceleration obtained from the accelerogram of that component. Horizontal PGAs have commonly been used to describe ground motions because of their natural relationship to inertial forces; indeed, the largest dynamic forces induced in certain types of structures (i.e., very stiff structures) are closely related to the PGA.

The *peak ground velocity* (PGV) is another useful parameter for characterization of ground motion amplitude. PGV represents the maximum absolute value of velocity at the ground. The velocity is less sensitive to the higher-frequency components of the ground motion, therefore PGV is more likely than the PGA to characterize ground motion amplitude accurately at intermediate frequencies.

The *peak ground displacement* (PGD) is the maximum absolute value of displacement at the ground and it is generally associated with the lower frequency components of an earthquake motion. PGD is less commonly used as IM respect to PGV and PGA due to signal processing errors in the filtering and integration of accelerograms and due to long-period noise.

Loads produced by earthquakes are complex and characterized by components of motion that span a broad range of frequencies. The frequency content of ground motion describes how the amplitude of a ground motion is distributed among different frequencies. Since the dynamic response of civil structures is very sensitive to the frequency at which they are loaded, alternative IMs based on the frequency content of strong motion have been proposed to characterize the ground motion. Among them, the most used in in earthquake engineering practice is the *response spectrum* defined as the maximum response (in terms of displacement, velocity or acceleration) of an elastic single degree- of-freedom (SDoF) system to a particular input motion as a function of the natural frequency (or natural period) and damping ratio of the SDoF system.

Duration of strong motion is also important for many physical processes that are sensitive to the number of load that occur during an earthquake (e.g. the degradation of stiffness and strength of certain old types of structures). The duration of strong ground motion is related to the time required for release of accumulated strain energy by rupture along the fault. As consequence, the duration increase with increasing earthquake magnitude. Due to the ground noise, a seismograph always receives little shakes that doesn't come from any seismic activity: wind, sea waves, road traffic and industrial activity are the principal causes for noise. As consequence some conventional measures of seismic duration have been introduced. Among them the *Bracketed duration* and the *Trifunac* and *Brady duration* are the most used. The *Bracketed duration* is defined as the time between the first and last exceedance of a threshold acceleration (usually 0.05 g). *Trifunac* and *Brady duration* depends on the amount of energy released in the site during the earthquake: it is the time interval between the 5% of the energy and the 95% of total released energy.

The preceding IMs are related primarily to only one of the three characteristics, i.e. the amplitude, frequency content, or duration of the ground motion. Since all of them are important, intensity measures that reflect more than one characteristic are very useful. An example is represented by the Arias Intensity. The *Arias Intensity* I_A is a descriptor of the energy of the earthquake and it is analytically defined as:

$$I_A = \frac{\pi}{2g} \int_0^\infty \ddot{u}_g(t) dt \tag{7}$$

where g is the acceleration due to gravity and $\ddot{u}_g(t)$ is the acceleration-time history in units of g.

I_A has the dimension of velocity. Since it is obtained by integration over the entire duration, its value is independent of the method used to define the duration of strong motion. This IM is well correlated to the damage level in soil and geotechnical structures.

2.6.1 Prediction of strong ground motion

Empirically based estimates of ground motion parameters are the oldest estimates in seismic hazard analysis, dating from the 1960s. Ground motion prediction equations (GMPEs) provide probabilistic distribution of the chosen IM (the predicted variable characterizing the level of shaking) conditional on a set of explanatory variables. GMPEs are obtained by regression of recorded data from historical events and they are popular for regions where many data are available. GMPEs may change with time as additional strong motion data become available. Most predictive relationships are updated in the literature every 3 to 5 years or after the occurrence of a major event. In stable continental regions where low seismicity rates allow to collect only limited data, theoretical approaches are generally used instead. In such cases, synthetic time histories are simulated in order to enlarge the strong motion database on which regression techniques can then be applied.

The most used explanatory variables are the magnitude, the source-to-site distance and coefficients to take into account for style of faulting, wave propagation path, and/or local site conditions. The typical form is expressed in the following equation:

$$\log IM = \overline{\log IM}(M, R, \underline{\theta}) + \varepsilon \tag{8}$$

$\overline{\log IM}(M, R, \underline{\theta})$ is the mean of the logs conditional on parameters such as magnitude (M), source-to-site distance (R), and others $\underline{\theta}$); the difference between the observed and the predicted ground motion is the ground motion residual that represents the unexplained part of the model.

The functional form is usually selected to reflect the mechanics of the ground motion process as closely as possible. Common forms are based on the following observations (Kramer, 1996):

- i Peak values of strong motion parameters are approximately log normally distributed.
- ii Earthquake magnitude is typically defined as the logarithm of recorded peak amplitude.
- iii Body wave ($P-$ and $S-$wave) amplitudes decrease according to 1/R and surface wave (primarily Rayleigh wave) amplitudes according to $1/\sqrt{R}$.
- iv The area over which fault rupture occurs increases with increasing earthquake magnitude. As a result, some of the waves that produce strong motion at a site arrive from a distance, R, and some arrive from greater distances. The effective distance, therefore, is greater than R by an amount that increases with increasing magnitude.
- v Some of the energy carried by stress waves is absorbed by the materials they travel through (material damping). This material damping causes ground friction amplitudes to decrease exponentially with R.
- vi Ground motion parameters may be influenced by source characteristics (e.g. type of fault) or site characteristics.

Therefore, a typical GMPE may have the following form:

$$\begin{aligned} \underbrace{\log IM}_{\text{i}} = \underbrace{C_1 + C_2 M + C_3 M^{C_4}}_{\text{ii}} + \underbrace{C_5 \log[R}_{\text{iii}} + \underbrace{C_6 \exp(C, M)}_{\text{iv}}]^{1/2} \\ + \underbrace{C_8 R}_{\text{v}} + \underbrace{f(source) + f(site)}_{\text{vi}} + \varepsilon \end{aligned} \tag{9}$$

The residual term ε is assumed normally distributed with zero mean and standard deviation $\sigma_{\log IM}$ that describes uncertainty in the value of the ground motion parameters given by the predictive relationship (all the effects which are not accounted for in the chosen functional form).

In the most recents GMPEs the residual is expressed as the sum of two components: an inter-event term, which is constant for each earthquake (common for all sites) and represents average source effects not explicitly appearing in the model covariates, and an intra-event term representing the site-to-site variability of the ground motion parameter. Therefore for a particular

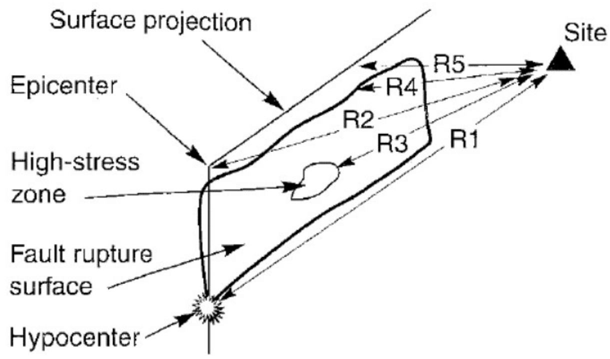

Figure 10: Distance metrics used in GMPEs (Kramer, 1996).

site p and earthquake j the logs of ground-motion intensities and related heterogeneity may be expressed as in the following equation:

$$\log IM_{pj} = \overline{\log IM_{pj}}(M, R, \underline{\theta}) + \eta_j + \varepsilon_{pj}, \tag{10}$$

where η_j denotes the inter-event residual, which is a constant term for all sites in a given earthquake and represents a systematic deviation from the mean of the specific seismic event (i.e. average source effects not explicitly appearing in the model covariates); and ε_{pj} is the intra-event variability of ground motion. Residuals ε_{pj} and η_j are usually assumed to be *independent random variables*, normally distributed with zero mean and standard deviation σ_{intra} and σ_{inter}, respectively. Then, $\log IM_{pj}$ is modeled as a normal random variable with mean $\overline{\log IM_{pj}}(M, R, \underline{\theta})$ and standard deviation σ_t, where $\sigma_t = \sqrt{\sigma^2_{intra} + \sigma^2_{inter}}$.

When using any GMPE, it is very important to know how parameters such as magnitude and source to site distance are defined. Concerning the distance, it is important to point out the different metrics used in the ground motion models (Figure 10): R1 and R2 are the hypocentral and epicentral distance, which are the easiest distance to determine after an earthquake. R3 is the distance to the zone of highest energy release. Since rupture of this zone is likely to produce the peak ground motion amplitudes, it represent the best distance measure for peak amplitude predictive relationships. Unfortunately, its location is difficult to determine after an earthquake and nearly impossible to predict before an earthquake. R4 is the closest distance to the zone of rupture and R5 is the closest distance to the surface projection of the fault rupture (also known as Joyner-Boore distance).

References

Cornell, C.A. (1968). Engineering seismic risk analysis, Bull. Seism. Soc. Am., 58, 1583-1606.

Grotzinger, J.P. and Jordan, T.H. (2014). Understanding Earth. W.H. Freeman and Company. New York, NY 10010.

Kramer, S.L. (1996) Geotechnical earthquake engineering. Prentice Hall, Upper Saddle River, N.J.

McGuire, R.K. (2004) Seismic Hazard and Risk Analysis, Earthquake Engineering Research Institute Publication, Report MNO-10, Oakland, CA, USA

SHAAC, Senior Seismic Hazard Analysis Committee. (1997) Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts, US Nuclear Regulatory Commission report CR-6372, Washington DC.

Wells, D.L. and Coppersmith, K.J. (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull. Seism. Soc. Am., 84, 974-1002.

PART II
Probabilistic Seismic Hazard Analysis

3 Probabilistic Hazard Model

Probabilistic Seismic Hazard Analysis (PSHA) evaluates the exceedance (or occurrence) probability of a given ground motion intensity measure threshold at given site and time interval.

This method has many applications in the field of earthquake engineering, including the design or retrofitting of critical facilities. More recently, results of PSHA have also been used for the determination of earthquake insurance coverage of private homes and businesses. PSHA provides a framework in which uncertainties are quantified, and combined in a rational manner to provide a comprehensive description of the seismic hazard. These uncertainties typically include magnitude size, earthquake location, soil condition, and rate of occurrence of earthquakes.

In detail, the calculation of seismic hazard is based on the Total Probability Theorem:

$$P[E] = \int_s P(E|S=s) f_S(s) ds, \tag{11}$$

where $P[E]$ represents the probability that the event E occurs, $P(E|S)$ is the conditional probability of the event E given the occurrence of the event S and $f_S(s)$ is the PDF of S, being S a continuous random variable.

In the context of seismic hazard, S is described by the magnitude M and the source to site distance R, and E is the event of overcoming an intensity measure level im at a given site. Given N_s seismic sources (e.g. multiple faults or multiple area sources), the (11) for a given source n becomes

$$P(IM > im) = \int_s P(IM > im|M = m, R = r) f^{(n)}_{R|M}(r|m) f^{(n)}_M(m) dr dm. \tag{12}$$

The (12) expresses the probability that a fixed value of ground motion im is exceeded at a given site, given the occurrence of *random* earthquake from the seismic source n. Observe that in case R is the epicentric or hypocentric distance R and M are statistically independent, i.e. $f^{(n)}_{R,M}(r,m) = f^{(n)}_R(r) f^{(n)}_M(m)$.

The general procedure is outlined by Cornell (1968) and it is based on four main steps:

i **Earthquake Source Characterization.** Identification and classification of the N_s sources. Each source can be represented as area source, fault source, or, rarely, point sources, depending upon the geological nature of the sources and available data, Section 3.1. Definition of $f^{(n)}_{r|m}(r|m)$ or $f^{(n)}_r(r)$ (in case $R \perp\!\!\!\perp M$). The distribution of $R|M$ or simply R depends on the definition of the distance the seismicity source, and wether considering or not the rupture size, Section 3.1.

ii **Earthquake size.** Definition of $f^{(n)}_M(m)$ for each source n, based on magnitude recurrence relationship, Section 3.2.

iii **Ground motion estimation.** Definition of $P(IM > im|M = m, R = r)$. These are empirical regression models named *ground motion prediction equations* (GMPE), Section 3.3. For a given magnitude and distance they define the probability of exceedance of a given intensity measure level im.

iv **Hazard computation.** Solution of the integral 12 for all the N_s sources. Section 3.4.

This procedure is based on the assumption that earthquakes form a stochastic process. Because the earthquake can be considered instantaneous event with respect to the time periods of engineering interest, counting processes are the stochastic way to model earthquake occurrence which is a fundamental term of PSHA. Most applications of PSHA are based on the assumption that the earthquake process is memoryless, that is, there is no memory of the time, size and location of preceding events. This assumption is typically made by defining the occurrence of earthquakes as Homogenous Poisson Process (HPP) characterized by an exponential distribution of earthquake recurrence intervals (further details about this assumption will be provided in the following).

3.1 Earthquake source characterization

The first step required for PSHA is represented by the identification and the characterization of seismic sources that can affect the site of interest. The identification of seismic source zones is based upon the interpretation of geological, geophysical, and seismological data.

There are two general types (McGuire, 2004):

i **Fault sources**: are faults or zones for which the tectonic features causing earthquakes have been identified. These are usually individual faults, but they may be zones comprising multiple faults or regions of faulting if surface evidence is lacking but the faults are suspected from evidence (e.g. seismicity patterns). Although originally only modeled as linear sources, most fault source models now have multi-planar features and ruptures are assumed to be distributed over the entire fault plane.

ii **Area sources**: areas within for which future seismicity is assumed to have distributions of source properties and locations of energy release that do not vary in time and space. These are regions defined by polygons within which the seismicity is assumed uniform in terms of type and distribution of earthquakes.

In the context of PSHA, sources may be similar to or somewhat different than, the actual source, depending on the relative geometry of the source and site of interest and on the quality of the information about the sources (Kramer, 1996). In Figure 11 three examples of different source geometries are shown. The relative short fault shown in Figure 11a can be modelled as a point source since the distance between any point along its length and the site of interest is nearly constant. The fault plane shown in Figure 11b is characterized by a depth that is sufficiently small that the fault can be simplified as a linear source. The available data of the source shown in Figure 11c are not sufficient to determine the actual geometry, so it represented as a volumetric source.

3.1.1 Distance

For a given earthquake source, it is generally assumed that earthquakes will occur with equal probability at any location on the source (i.e. uniformly distributed hypocenter location). Given that locations are uniformly distributed, the geometry of the source is used to identify the probability distribution of source-to-site distances.

Note that several distance definitions are used in PSHA (See Section 2.6.1 and Figure 10). Epicentral and hypocentral distances only consider the location of the rupture initiation; some others instead need to account for the fact that ruptures occur over a plane rather than a single point in space. The choice of distance definition will depend upon the required input to the GMPE. In case the estimate of fault rupture dimensions is needed, two methods can be applied: estimating the dimensions based directly on the size of the fault rupture plane or by basing the estimate on the size of the aftershock zone.

In the following examples we will derive the analytical distance distributions for simple geometric sources. Note that only the epicentral distance is considered here for simplicity. In the more general cases, (e.g.

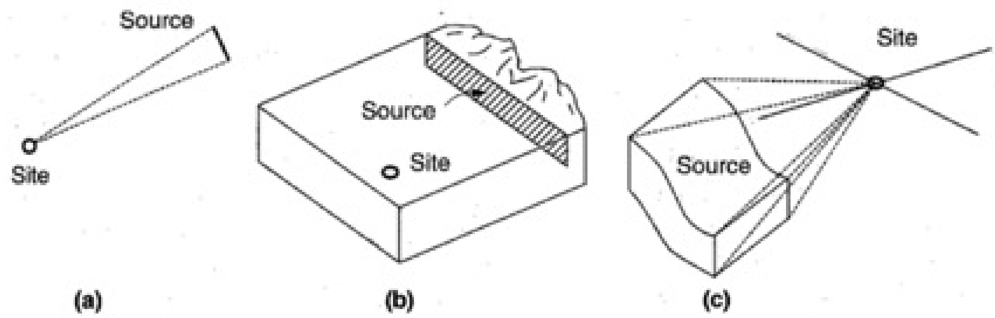

Figure 11: Example of different source geometries (Kramer, 1996).

complex geometries) distance distributions are more complicate to derive analytically and sometimes it is required numerical simulations.

Example III-a. Point source.

The problem is completely deterministic since the distance r is fixed, Figure 12-a). Then, $f_R(r') = \delta(r - r')$.

Example III-b. Linear source.

For the linear source in Figure 12-b), the probability that an earthquake occurs on the small segment of the fault between $L = l$ and $L = l + dl$ is the same as the probability that it occurs between $R = r$ and $R = r + dr$, i.e. $f(l)dl = f(r)dr$.

As $l^2 = r^2 - r_{min}^2$, and considering that the earthquakes are assumed to uniformly distributed over the length of the fault L_f, $f(l) = l/L_f$, the CDF of R is

$$F(r) = P(R \leq r) = \frac{\text{length of the fault within distance } r}{\text{total length of the fault}} \tag{13}$$

which can be written as

$$\begin{aligned} F(r) &= 0, \text{ for } r < r_{min} \\ &= \frac{\sqrt{r^2 - r_{min}^2}}{L_f}, \text{ for } r_{min} \leq r \leq r_{max} \\ &= 1 \text{ for } r > r_{max}. \end{aligned} \tag{14}$$

The PDF is obtained by differentiation of the (14) as

$$\begin{aligned} f(r) &= 0, \text{ for } r < r_{min} \\ &= \frac{r}{L_f \sqrt{r^2 - r_{min}^2}}, \text{ for } r_{min} \leq r \leq r_{max} \\ &= 0 \text{ for } r > r_{max}. \end{aligned} \tag{15}$$

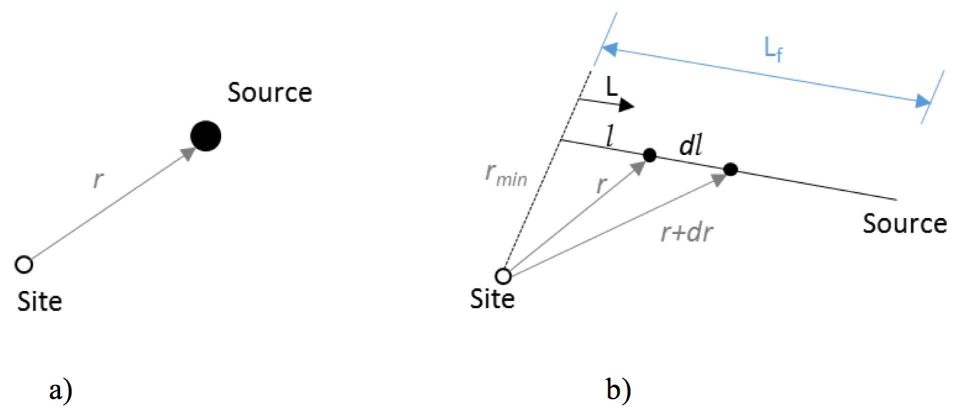

Figure 12: Example of a point source (a) and a linear source (b).

Example III-b. Area source.

This example is also the solution of the Problem *iii* in Section 3.1.2 of Fundamental of Probabilities. The site is located in an area source shown in Figure 13. Earthquakes are equally likely to occur anywhere in the area within K km from the site.

The probabilistic distribution of an epicenter being located at a distance of less than r is then equal to

$$F(r) = P(R \leq r) = \frac{\text{area of the circle with radius } r}{\text{area of the circle with radius } K} \tag{16}$$

which can be written as

$$\begin{aligned} F(r) &= \frac{r^2}{K^2}, \text{ for } 0 < r \leq K \\ &= 1 \text{ for } r \geq K. \end{aligned} \tag{17}$$

The PDF is obtained by differentiation of the (17) as

$$\begin{aligned} f(r) &= \frac{r}{2K^2}, \text{ for } 0 \leq r < K \\ &= 0 \text{ for } r \geq k. \end{aligned} \tag{18}$$

3.2 Earthquake size

Once the source characterization is completed, we define the distribution of the $f_M(m)$. This is usually based on a recurrence model describing *"the chance of an earthquake of a given size occurring anywhere inside the source during a specified period of time"*. A basic assumption of PSHA is that the recurrence law obtained from past seismicity is appropriate for the prediction of future earthquakes.

In many applications the exponential probability distribution is used to represent the relative frequency of different earthquake magnitudes(McGuire, 2004). In particular, the most used recurrence law model is the one proposed by Gutenberg and Richter (1954). The *Gutenberg-Richter law (G-R law)* expresses the relationship between the magnitude and rate of cumulative number of earthquakes in any given region:

$$\log \lambda(m) = a - bm, \tag{19}$$

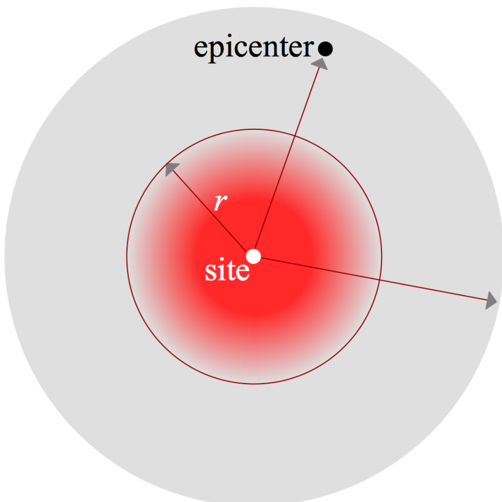

Figure 13: Example of an area source.

where log is the logarithm base 10, $\lambda(m)$ is the mean annual rate of exceedance of magnitude m, and a and b are constants. The a value indicates the overall rate of earthquakes of the source, and the b value indicates the relative ratio of small and large magnitudes (a typical value of b is 1). As b increases, the number of larger magnitude events decreases compared to those of smaller magnitudes. Equation (20) is also expressed as

$$\lambda(m) = 10^{a-bm} = \lambda_0 \exp(-\beta m) = \exp(\alpha - \beta m), \tag{20}$$

where $\lambda_0 = 10^a$ represents the mean yearly number of earthquakes with $m \geq 0$, $\alpha = \ln(10)a$ and $\beta = \ln(10)b$.

The general formulation of the G-R law covers an infinite range of magnitudes, but often a lower and an upper-bound, m_{min} and m_{max} respectively, are used resulting in a truncated exponential distribution for magnitude frequency, named bounded G-R law (14).

The minimum magnitude is generally linked to the minimum magnitude which is believed to produce damages to the structures. It is usually set at values from about 4.0 to 5.0.

The truncation at m_{max} may arise because the magnitude scale saturates (see Section 2.5) or because the seismic zone cannot physically generate magnitudes above this value. It follows that the selection of m_{max} is complicate since it is generally estimated by using geologic evidence, geophysical data, analogies to similar tectonic regimes, or other methods (McGuire, 2004). The m_{max} can be estimated in several ways; the simplest is to take the maximum value historically observed and conservatively slightly increase this value, for example by 0.2 units. More advanced methods estimate m_{max} from the earthquake catalogue of the source considered or from the geometry of the source (e.g. Wells and Coppersmith, 1994, Table 1).

The CDF of the bounded GR-law considering only the lower bound can be obtained as follow

$$\begin{aligned} F_M(m) &= P(M \leq m | M > m_{min}) \\ &= \frac{\lambda_{min} - \lambda(m)}{\lambda_{min}}, \\ &= \frac{10^{a-bm_{min}} - 10^{a-bm}}{10^{a-bm_{min}}} \\ &= 1 - 10^{-b(m-m_{min})} \\ &= 1 - \exp[-\beta(m - m_{min})], \; m > m_{min} \end{aligned} \tag{21}$$

and the PDF by differentiation as

$$\begin{aligned} f(m) &= \frac{d}{dm} F(m) \\ &= b\ln(10)10^{-b}(m - m_{min}) \qquad (22) \\ &= \beta \exp[-\beta(m - m_{min})], \; m > m_{m_{min}}. \qquad (23) \end{aligned}$$

It follows that the mean annual rate can be written as

$$\lambda(m) = \lambda_{min} 10^{-b(m-m_{min})} = \lambda_{min} \exp[-\beta(m - m_{min})], \tag{24}$$

where $\lambda_{min} = 10^{a-b_{min}} = \exp(\alpha - \beta m_{min})$ represents the mean yearly number of earthquakes with $m \geq m_{min}$. If the upper bound is also considered, then the (21) becomes

$$\begin{aligned} F_M(m) &= \frac{1 - 10^{-b(m-m_{min})}}{1 - 10^{-b(m_{max}-m_{min})}}, \qquad (25) \\ &= \frac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]}, \; m_{min} \leq m \leq m_{max}, \qquad (26) \end{aligned}$$

with PDF as

$$\begin{aligned} f_M(m) &= \frac{b\ln(10)10^{-b(m-m_{min})}}{1 - 10^{-b(m_{max}-m_{min})}}, \qquad (27) \\ &= \frac{\beta \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]}, \; m_{min} \leq m \leq m_{max}, \qquad (28) \end{aligned}$$

and mean annual rate of exceedance as

$$\lambda(m) = \lambda_{min} \frac{10^{-b(m-m_{min})} - 10^{-b(m_{max}-m_{min})}}{1 - 10^{-b(m_{max}-m_{min})}}, \tag{29}$$

$$= \lambda_{min} \frac{\exp[-\beta(m - m_{min})] - \exp[-\beta(m_{max} - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]}, \; m_{min} \leq m \leq m_{max}. \tag{30}$$

Important considerations

- The use of the G-R law is not restricted to the use of magnitude as descriptor of the earthquake size. As explained in Section 2.5, intensity has also been used.
- Given a seismic source, the constants a and b (equation (20)) are usually estimated using statistical analysis of historical data (preinstrumental and instrumental seismicity) with additional constraining data provided by geological evidence. Historical data are usually provided by seismic catalogues, which summarized all important information about past events: e.g., date, hypocenter and epicenter location, earthquake size, fault mechanism, etc. The use of both instrumental and preinstrumental events implies a seismic catalogue that may contain both magnitude (based on different scales, see Section 2.5) and intensity data. Then, there is a need for conversion between the measures. This is expressed as: first by statistical and approximated correlation between macroseismic damage measures (i.e. intensity scales); second by a quantitative classification of earthquake intensity and between different magnitude scales. As an example, here it is reported a relationship between surface magnitude and epicentral intensity (a qualitative scale for damage estimation at the epicentral area) estimated for the Parametric Catalogue of Italian Earthquakes (Catalogo Parametrico dei Terremoti Italiani - CPTI): $M_s = 0.561 I_0 + 0.937$. Note that whatever magnitude scale is chosen for the conversion, it must be consistent with the method chosen for estimating earthquake ground motion (See Section 2.6 and 3.4).
- Another important aspect to consider for the estimation of the G-R parameters is the *completeness* of the seismic catalogue in terms of intensity/magnitude and time intervals in which a certain intensity/magnitude range is likely to be completely reported (i.e. it is possible to assume that all the earthquakes of the considered magnitude are recorded in the catalogue). Geometry and

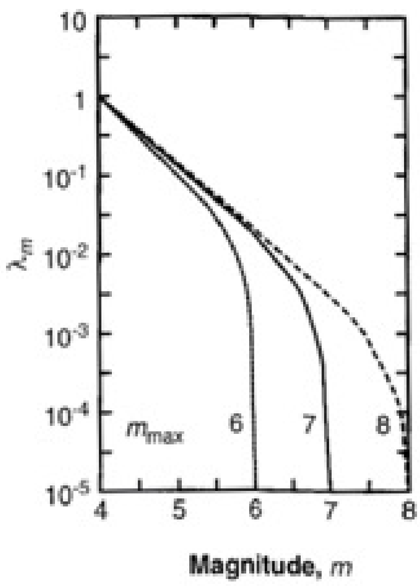

Figure 14: Bounded G-R law considering $m_{min} = 4$ and different m_{max} (Kramer, 1996).

coverage of the seismic network, or malfunction of seismic stations are the main responsible of the incompleteness of instrumental data. Analysis of catalogue completeness can be done via several methods such expert analysis of historical data and sources or visual procedures.

- Records of historical seismicity may be distorted by the presence of dependents events represented by the group of foreshock and aftershock which occur respectively before and after the mainshock in a variable time and space window. Since the PSHA is aimed at evaluating the seismic hazard associated to the main events (i.e. the mainshock), dependent events must be removed from the seismic catalogue and their effects accounted in separate analyses. One of the first analytical model for removing group of foreshock and aftershock was proposed by Gardner and Knopoff (1974) allowing the identification of spatial and temporal window of clustering around the mainshock as a function of mainshock magnitude.

3.2.1 Alternative magnitude distributions

Several magnitude distributions other that the truncated exponential distribution are available for modeling earthquake magnitudes as the Characteristic Magnitude Distribution. This distribution is used when continuous distributions encompassing all magnitudes are not appropriate (McGuire, 2004). Geological data may indicate that the characteristic earthquakes occur more frequently that would be implied by extrapolation of the GR-law from high exceedance rate (low magnitude) to low exceedance rates (high magnitude). This may be the case of individual sources (i.e. single faults) that usually generate earthquakes of similar size (i.e. characteristic earthquakes) at or near their maximum magnitude. The result is a more complex recurrence law that is governed by historical seismicity data at low magnitudes and geological data at high magnitudes (Kramer, 1996).

In most of the cases a continuous exponential distribution may be adequate for events up to, say, $m = 6-7$. Large earthquakes (e.g. $m = 7.5 - 8$) may occur with a characteristic magnitude whose frequency of occurrence is higher than obtained extrapolating from the smaller magnitude earthquakes. In this case, two truncated exponential distributions could model the magnitude occurrence of future events: a distribution between $m_{min} = 5$ and $m_{max} = 7$ with a standard β value (e.g. $\beta = 2.3$), and a separate distribution between $m_{min} = 7.5$ and $m_{max} = 8$ with $\beta = 0$ representing the equal likelihood of a characteristic event magnitude in that range. Seismicity data can be used then to estimate the rate for the first distribution while the rate for characteristic events can be estimated with geological evidence.

One form of the characteristic magnitude distribution is illustrated in Figure 15 developed by Young and Coppersmith (1985). This generalized recurrence law combines an exponential magnitude distribution at lower magnitudes with an uniform distribution in the proximity of the characteristic earthquake. The figure also shows a comparison with the bounded G-R law assuming the same upper magnitude bound value and slip rate.

Important considerations

- To adopt an alternative magnitude distribution the only requirement is to use the appropriate PDF. All the other steps of the PSHA remain identical.
- Evaluation of which model is most appropriate for a given source is hampered by brevity of historical and /or instrumental records. The seismicity records of the last decades for the major sources of the southern California suggest that while the available data were not sufficient to disprove the G-R model, the characteristic model better represented the observed distribution of magnitudes.

3.3 Ground motion estimation

Once the distribution of potential earthquakes magnitudes and locations has been identified, we can evaluate the ground motion at the site. There are two basic steps to estimate the ground motion at the site.

i Identification of the important characteristics of the ground motion (i.e. the intensity measures, see Section 2.6). The selection of the IM of interest (amplitude, frequency and/or duration-based) depends on the element at risk under consideration, the effects being considered in the analysis and approach that is followed for the derivation of fragility curves. Several concepts and quantities are commonly used to assist in identifying the optimum intensity measure for the required purpose. These are defined as practicality, effectiveness, efficiency, sufficiency and robustness (Cornell et al., 2002).

ii Estimation of the probability distribution of the selected IM must be estimated as a function of predictor variables such as the earthquake source properties, the relative location of the earthquake respect to the site, and the soil conditions.

Ground motion prediction equations (GMPEs) are usually adopted to evaluate the probability that a particular IM exceeds a certain value, im, for a given earthquake $M = m$, occurring at a given distance, $R = r$, (as illustrated graphically in Figure 16). In probabilistic terms, this is written as $P[IM > im|R = r, M = m] = 1 - F_{IM|RM}(im|r, m)$.

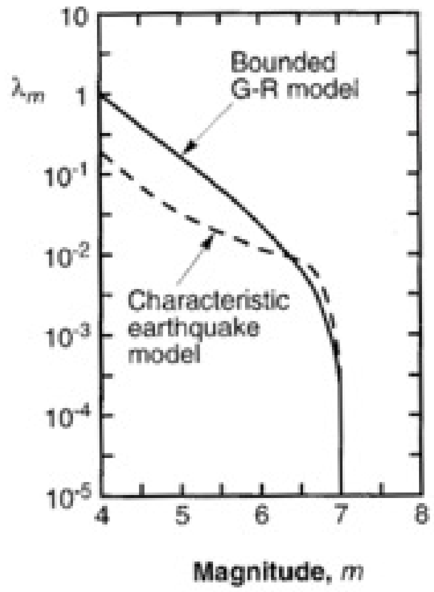

Figure 15: Comparison of recurrence laws from Characteristic and Bounded G-R magnitude models (Kramer, 1996).

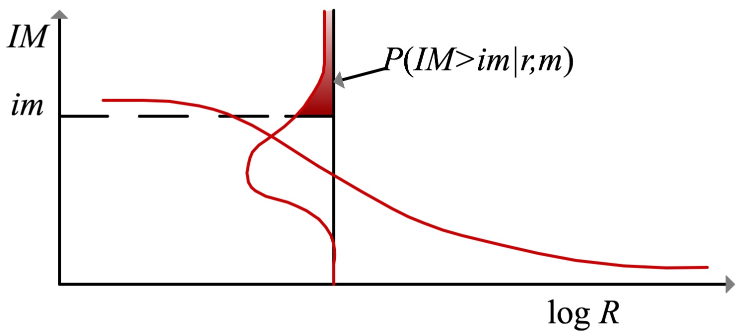

Figure 16: Schematic illustration of conditional probability of exceeding a particular value of a ground motion parameter for a given magnitude and distance.

In general, the conditional distribution of the ground motion intensity measure, i.e. $F(im|r, m)$ is assumed log normal. It follow that the logarithm of the conditional intensity measure is normally distributed. Then,

$$P[IM > im|R = r, M = m] = \int_{im}^{\infty} f_{IM|RM}(u)du = \frac{1}{\sigma_T\sqrt{2\pi}} \int_{im}^{\infty} \exp\left[-\frac{(\ln u - \overline{\ln IM})^2}{2\sigma_T^2}\right] du \tag{31}$$

where $\overline{\log IM}$ and σ_T are the mean value and the standard deviation of $\log Y$(See Section 2.6.1).

Important considerations

- Over decades of development, the prediction models have become complex, consisting of many parameters (see Section 2.6.1). More specifically, many factors related to source and site characteristics are considered. One of the most fundamental is the tectonic region in which the source is located. Typically, GMPEs are developed independently for each region (active, subduction and stable continental zones). Another important factor is the faulting mechanism that influences in particular amplitude and frequency content. Effects of local site conditions are represented in many forms, ranging from a simple constant to more complex terms that try to characterize non-linearity in ground response. Many recent works analyzed also the importance of near-fault effects (typically defined within 20-60 km of fault rupture), concluding that ground motion at near source cites are sensitive to the effect of the "rupture directivity" that affect duration and long period energy.
- As mentioned in Section 2.6, empirical or theoretical methods can be applied to estimate the ground motion. Stochastic methods are used to generate ground motions when there is insufficient amount of ground motion recordings to develop empirically-based equations. These methods are commonly used in stable tectonic regions and for high frequency motions characterized by a large magnitude and short source-to-site distance. Standard stochastic methods are based on the assumption that the far-field shear wave energy generated by an earthquake source can be represented as a band-limited white Gaussian noise random process.

3.4 Hazard computation

3.4.1 Seismic hazard curve

The seismic hazard curve is a function representing the annual frequency of exceeding various levels of ground shaking (i.e. the IM) at a specific site. The curve is obtained by integration of the first three steps over all possible magnitudes and earthquakes locations.

Seismic hazard curves are obtained for individual sources and, then, combined to express the aggregate hazard at a particular site. The aggregation is given after selecting a stochastic model for the occurrence of earthquakes in time. It is common to use the HPP model (Section 6.2 on Fundamentals of Probability and the following Section 3.4.2). For each sources the rate of exceedance of a given im is derived as Possion with random selection (Section 6.2.1 Fundamentals of Probability). In particular, given a rate of minimum earthquake of a source n, i.e. $\lambda_{min}^{(n)}$, the rate of of the event $IM > im$, $\lambda(im)$, is simply given by $\lambda^{(n)}(im) = \lambda_{min}^{(n)} P(IM > im|M \geq m_{min})$, where $P(IM > im|M \geq m_{min})$ is obtained by (12). Observe that here, for the sake of simplicity in the notation, $P(IM > im|M \geq m_{min})$ is equivalent to $P(IM > im)$.

Given N_s statistically independent sources, the aggregate hazard curve is derived by merging all the N_s Poisson process. Merging N_s statistically independent Poisson processes (Section 6.2.2 Fundamentals of Probability) is straightforward since it involves only the sum of the N_s different rates. Then, we can write

$$\lambda(im) = \sum_{n=1}^{N_s} \lambda_{min}^{(n)} \left[\int_{r_{min}}^{r_{max}} \int_{m_{min}}^{m_{max}} P(IM > im|M = m, R = r) f_{R|M}^{(n)}(r|m) f_M^{(n)}(m) dr dm\right], \tag{32}$$

or in case of $M \perp\!\!\!\perp R$

$$\lambda(im) = \sum_{n=1}^{N_s} \lambda_{min}^{(n)} \left[\int_{r_{min}}^{r_{max}} \int_{m_{min}}^{m_{max}} P(IM > im | M = m, R = r) f_R^{(n)}(r) f_M^{(n)}(m) dr dm \right]. \tag{33}$$

As the individual components of equations (32) and (33) are complicated, the integral can be evaluated analytically only for some simple situations (Cornell, 1968). Then, numerical integration is required. Among a variety of different techniques, the most simple and used approach is to divide the possible ranges of magnitude and distance into N_m and N_r segments, respectively. It follow that (33), for example, can be estimated as

$$\begin{aligned} \lambda(im) &\approx \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \sum_{l=1}^{N_m} \lambda_{min}^{(n)} P(IM > im | M^{(n)} = m_l, R^{(n)} = r_m) f_R^{(n)}(r_m) f_M^{(n)}(m_l) \Delta r \Delta m, && (34) \\ &\approx \sum_{n=1}^{N_s} \sum_{m=1}^{N_r} \sum_{l=1}^{N_m} \lambda_{min}^{(n)} P(IM > im | M^{(n)} = m_l, R^{(n)} = r_m) P(M^{(n)} = m_l) P(R^{(n)} = r_l). && (35) \end{aligned}$$

3.4.2 Earthquake Occurrence Model

Most applications of PSHA are based on the assumption that the earthquake process is memory-less, that is, there is no memory of the time, size and location of preceding events. Moreover, earthquake can be considered instantaneous event with respect to the time periods of engineering interest. We have seen that these assumptions are typically encompassed in a HPP process with rate λ (number of events in the unit time, that in the ordinary seismic application is the year) and exponential inter-arrival intervals (Section 6.2 on Fundamentals of Probability).

In particular, in the context of seismic hazard, HPPs possess the following properties:

i The probability of more than one occurrence during a very short time interval is negligible.

ii The number of earthquakes which occur in disjoint time intervals are independent (independent increments).

iii The number of earthquake events in any interval of length t is Poisson distributed, $P(N(t+s) - N(s) = n) = \frac{[\lambda(m)t]^n}{n!} \exp[-\lambda(m)t]$, with mean equal to $\lambda(m)t$ (stationary increments), where $\lambda(m)$ is the rate of earthquake with magnitude greater than m.

We have seen that the exponential inter-arrival time distribution has mean equal to $E[T] = 1/\lambda(m)$. This physically represents the mean time in which the next event is expected. Given this, the mean of the exponential is also defined as return period, i.e. $T_R = 1/\lambda(m)$. As a consequence of the memoryless property, T_R is a constant value during the whole observed time interval (50 years or one day after a seismic event, Poisson T_R is always equal to $1/\lambda(m)$). In other words all the information about the history of past earthquakes has no influence on the distribution of expected future events.

In traditional PSHA, the occurrence probability of *at least one* event in the observed time interval is of concern. It can be computed as:

$$P(N \geq 1, t) = 1 - P(N = 0, t) = 1 - \exp[-\lambda(m)t] \tag{36}$$

Thus relationship between occurrence probability of at least one event and return period is:

$$\begin{aligned} P(N \geq 1, t) &= 1 - \exp\left(-\frac{1}{T_R} t\right) \rightarrow \\ T_R &= -\frac{t}{\ln[1 - P(N \geq 1)]}. && (37) \end{aligned}$$

For example, the 10% occurrence probability of at least an event in a considered period of $t = 50$ [years], corresponds to a return period of $T_R \approx 475$[years].

Once the Poisson process is assumed, earthquake occurrence is completely defined by the mean annual number of earthquake occurrence $\lambda(m)$. As mentioned in the previous section, this allows two operations

i Poisson with random selection (known also as "filtering a Poisson process"). Section 6.2.1 Fundamentals of Probability

ii Combining Poisson processes (known also as "summing or merging Poisson processes"). Section 6.2.2 Fundamentals of Probability

Given the seismic source n the first operation is used to compute the rate of $IM > im$, i.e.

$$\begin{aligned} \lambda^{(n)}(im) &= \lambda_{min}^{(n)} P(IM > im|M > m_{min}) = \lambda_{min}^{(n)} P(IM > im) \quad (38) \\ &= \lambda_{min}^{(n)} \left[\int_{r_{min}}^{r_{max}} \int_{m_{min}}^{m_{max}} P(IM > im|M = m, R = r) f_{R|M}^{(n)}(r|m) f_M^{(n)}(m) dr dm \right] \quad (39) \end{aligned}$$

The second operation is used to aggregate all the N_s sources, i.e.

$$\lambda(im) = \sum_{n=1}^{N_s} \lambda_{min}^{(n)} P(IM > im) \quad (40)$$

to obtain either the (32) or the (33).

Finally, the aggregate probability of at least one exceedance of im in a period t years is equal to:

$$P(\text{at leas one event } IM > im, t) = F_t(im) = 1 - \exp[-\lambda(im)t]. \quad (41)$$

Given a time $t = 1$[year] and a specific site, the (41) represents the annual exceedance probability for earthquake of $IM > im$. The (41) is known also as the hazard curve. In the case $t = 1$[year] and for small value of λ_{im} the frequencies can be regarded as probabilities, since $1 - \exp[-\lambda(im)t] \approx \lambda(im)$.

Introducing T_R, we obtain

$$\begin{aligned} F_t(im) &= 1 - \exp[-\frac{1}{T_R} t] \rightarrow \\ T_R &= -\frac{t}{\ln[1 - F_t(im)]}. \quad (42) \end{aligned}$$

3.5 Uncertainty in PSHA

The definition and treatment of uncertainties represents an important aspect of PSHA. This involves identifying the intrinsic randomness of the modelled phenomenon, defined as *aleatory uncertainty* as well as the uncertainty related to our limited knowledge and data of the modelled phenomenon, defined as *epistemic uncertainty.*

The distinction between these two types of uncertainties is fundamental to understand where uncertainty originates and how it may be appropriately handled in hazard calculations.

The aleatory variability in a seismic hazard analysis is included directly in the calculations, specifically through the probability distributions of the parameters, and thus it directly influences the results. Epistemic uncertainty is related to the subjective decisions that are made as part of the process of carrying out the analysis, e.g:

i The definition of the geometry of the sources.

ii The assumptions regarding the completeness of the seismic catalogue

iii The definition of m_{max}

iv The selection of the GMPE

Since epistemic uncertainty is characterized by the use of alternative models (i.e. alternative probability distributions) it is not considered directly in the hazard calculations but rather is treated by developing alternative hazard curves.

A common way to handle epistemic uncertainty is through the use of *Logic Trees*. This methods allows the use of different models, each of which is assigned a weighting factor that represents the likelihood of that model being correct. Representing current scientific judgment on the merit of the alternative models, the weights are based on data collected from analogous regions, simplified physical models, and empirical observations. The use of the logic tree is to calculate the hazard following every brunch. The result of each analysis is finally weighted by the relative likelihood of its combination of branches, taken as the product of all the weightings along all the branches.

The logic tree shown in Figure 17 was used in the construction of the seismic hazard map of Italy (Gruppo di Lavoro, 2004). It considers the uncertainties related to the completeness of the catalogue of earthquakes, the method for calculating seismicity rates, and to the attenuation pattern. All these alternative options lead to a logic tree with 2x2x4 = 16 branches.

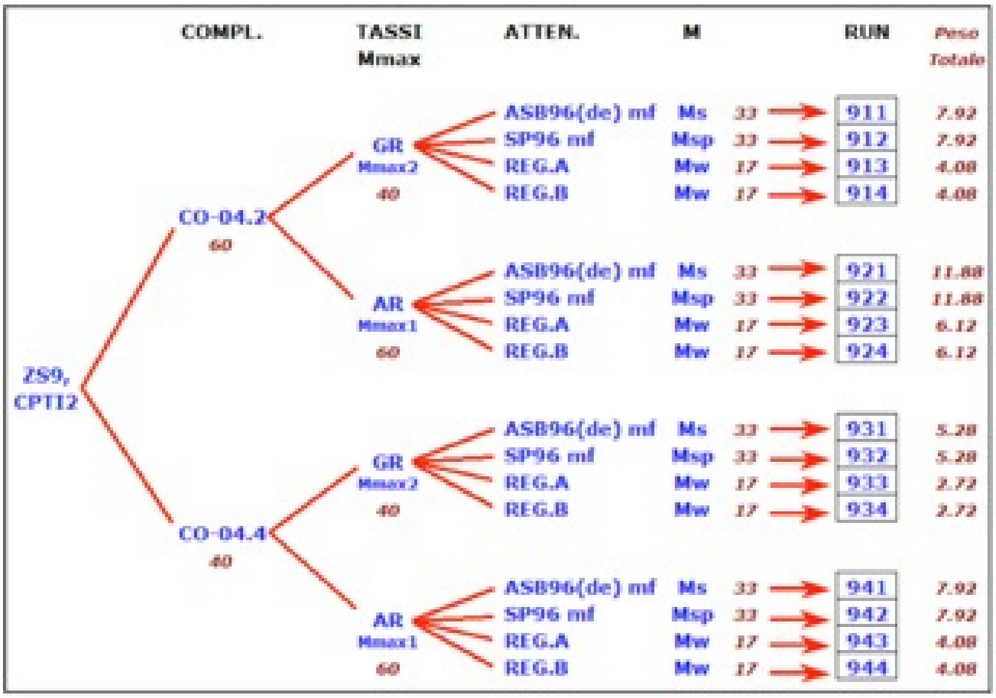

Figure 17: The logic tree used for the computation of the Italian seismic hazard map (Gruppo di Lavoro, 2004).

References

Baker J. W. (2008). An Introduction to Probabilistic Seismic Hazard Analysis (PSHA), White Paper, Version 1.3, 72 pp.

Cornell, C.A. (1968). Engineering seismic risk analysis, Bull. Seism. Soc. Am., 58, 1583-1606.

Cornell CA, Jalayer F, Hamburger RO, Foutch DA (2002) Probabilistic basis for 2000 SAC/FEMA steel moment frame guidelines. J. Struct. Eng., 1284, 26?533.

Gardner J.K. and Knopoff, L. (1974) Is the sequence of earthquakes in southern California, with aftershocks removed, Poissonian? Bull. Seism. Soc. Am., 64, 1363-1367.

Gruppo di Lavoro (2004) Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003. Rapporto conclusivo per il Dipartimento della Protezione Civile, INGV, Milano - Roma, 65 pp. + 5 App.

Gutenberg, B. and Richter C.F.(1942) Earthquake Magnitude, Intensity, Energy and Acceleration, Bull. Seism. Soc. Am., 32, 163-191.

Kramer, S.L. (1996) Geotechnical earthquake engineering. Prentice Hall, Upper Saddle River, N.J.

McGuire, R.K. (2004) Seismic Hazard and Risk Analysis, Earthquake Engineering Research Institute Publication, Report MNO-10, Oakland, CA, USA

Wells, D.L. and Coppersmith, K.J. (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull. Seism. Soc. Am., 84, 974-1002.

Youngs R.R. and Coppersmith K.J.(1985) Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard assessments. Bull. Seism. Soc. Am., 75, 939-964.