Good afternoon to everyone.

I’m Chiara Nardin, a civil engineer from the University of Trento, involved in a Ph.D. programme entitled Risk analysis of industrial plants based on Markov's processes.

Today, in these 2-hours lectures, we will discuss about PSHA and fragility analysis: by applying notions we are going to unfold to some already prepared MATLAB tutorials.

Before starting, all the materials presented are uploaded and hopefully updated on the GitHub link here below. Please do not hesitate to fork or to download and feel free to contact me for any issues.

-----------------------------------------------------------------------------------------------------------------------------------------------

Here is presented the course outline.

First, a brief introduction on PBEE and especially on the proposed framework of the PEER, i.e. Pacific Earthquake Research Center of Berkeley, here adopted.

Then, Probabilistic seismic hazard model is presented: starting from main concepts and analytical formulations to an implementation on MATLAB and a comparison of results with the ones provided by the Italian National Institute INGV {AINGVI}.

In the second hour, we will unfold main ideas on fragility concepts and discuss a case study implemented on MATLAB in order to perform fragility curves.

Finally, some references are provided.

-----------------------------------------------------------------------------------------------------------------------------------------------

Let’s start with PBEE.

Performance-Based Earthquake Engineering (PBEE) is a probabilistic framework, which supports designers and stakeholders in assessing design, evaluation, and planning of civil systems. The utmost goal of the PBEE is to replace the classical load-and-resistance-factor design, based on evaluation of failure probabilities of single components, by developing instead a framework capable of assessing performances of a system, with respect to different decision variables such as downtime, monetary losses and fatalities.

A schematic view of the PBEE framework is given in Figure 1.

Two axes define the PBEE framework, a discrete set of system performance objectives, that includes for example fully operational, operational, life safe and near collapse, and a set (discrete or continuous) of seismic hazard level, divided according to return period or exceedance probability.

The generated domain is therefore partitioned into an acceptable and not acceptable subdomain: here, in the former, designer and stakeholders identify the desired or optimal combination between performance objectives and hazard level in order to evaluate, according to its likelihood, the monetary cost of the system.

-----------------------------------------------------------------------------------------------------------------------------------------------

There are several alternatives and approaches to PBEE; anyway, in this class, we focus on the one provided by PEER.

The framework is simply a statement of the total probability theorem for the yearly mean number of events of a selected decision variable. You can see its analytical expression in Eq.(1): the so-called famous triple integral of PEER.

The underlying assumptions are here briefly reported and are basically: i) markovian structure, i.e. each of the conditional probabilities we are going to see can be read as statistical independent from the previous terms; ii) no aging effects are considered in the structure; iii) seismic events as a poissonian processes.

-----------------------------------------------------------------------------------------------------------------------------------------------

However, what is more important to understand reguards the structure of this formulation.

In fact, Eq. (1) is build on four distinct stages that can be considered and evaluated alone.

The first stage/step reguards

1. hazard analysis which evaluates the mean annual rate of exceedance of the IM, at the facility site.
2. structural/fragility analysis: the engineer creates a structural model of the system, where the input is derived from the hazard analysis and the output is computed via structural analysis.
3. damage analysis: evaluates the relationship between EDP(i.e. the output of fragility analysis) and DM damage measure
4. loss analysis: (the yellow one) decision variable measures the seismic performance of the facility in terms of the stakeholder interest

%each of these steps can be addressed by a different research group.

-----------------------------------------------------------------------------------------------------------------------------------------------

Here’s who is who. But what is again more important to understand is that each of this arrows stands for a model analysis that is statistical independent from the other analysis.

-----------------------------------------------------------------------------------------------------------------------------------------------

So, the first step of PBEE is SHA, that represent the basis for many earthquake design codes practices and for seismic risk analysis of civil systems. Reliable estimates of a seismic hazard are paramounts to minimize loss of life, property damage and social and economic disruption.

-----------------------------------------------------------------------------------------------------------------------------------------------

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.* So, 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 efiects. However, it does not provide any information regarding the likelihood of occurrence of the controlling earthquakes.

Probabilistic seismic hazard analysis instead (PSHA) provides an uncertainty quantification framework

to address the hazard analysis.

Here we will focus our attention on this probabilistic approach.

-----------------------------------------------------------------------------------------------------------------------------------------------

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.

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 as reported in eq. (1)

The (1) 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*.

-----------------------------------------------------------------------------------------------------------------------------------------------

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

* source characterization: Identification and classification of the Ns 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

* earthquake size: Definition of the PDF magnitude for each source n, based onthe magnitude recurrence relationship or law, i.e. for example the Gutenberg-Richter one;

Ground motion estimation. Definition of P(IM > im|M = m;R = r). conditional probabilities, or For a given

magnitude and distance the definition of the probability of exceedance of a given intensity measure level

im.

* Hazard computation. Solution of the the triple integral for all the Ns sources.

-----------------------------------------------------------------------------------------------------------------------------------------------

So, here the main assumptions underlying PSHA are remarked:

* Earthquakes consists in a stochastic process; and can be considered instantaneous and memoryless events

These let us make a reasonable assumption in define the occurrence of ground motions as homogeneus poisson process and that means that the equation describing the recurrence distribution is an exponential one, as depicted aside.

-----------------------------------------------------------------------------------------------------------------------------------------------

Let’s enter in the detailed steps of the PSHA.

The first goal is to identify the nature of the sources: that could be fault both individual or multiple, but also area sources define by polygons in which seismicity is assumed uniform. The identification is based upon the interpretation of geological and seismological data. So, identify but also characterize the seismic source in terms also of distances, or in other term to characterize the geometry of the source so that is possible to identify the probability distribution of source-to-site distances.

-----------------------------------------------------------------------------------------------------------------------------------------------

Once the source characterization is completed, we define the distribution of the fM(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”.

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.

where log is the logarithm base 10, \_(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).

-----------------------------------------------------------------------------------------------------------------------------------------------

Some observations are needed. The general formulation of the G-R law covers an ideally infinite range of magnitudes, but often a lower and an upper-bound, mmin and mmax 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 mmax may arise because the magnitude scale saturates or better because the seismic zone cannot physically generate magnitudes above this value.

Then another point, regards the constants *a* and b that are usually estimated using statistical analysis of historical data (preinstrumental and instrumental seismicity) with additional constraining data provided by geological evidence.

Finally Another important aspect to consider for the estimation of the G-R parameters are the completeness and undistortion 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.

-----------------------------------------------------------------------------------------------------------------------------------------------

Here some picture of the distribution of seismic events given magnitude are represent and plotted. IN particular it could be appreciated the comparison or better the application of both GR and GR bounded.

-----------------------------------------------------------------------------------------------------------------------------------------------

Once the distribution of potential earthquakes magnitudes and locations has been identi\_ed, we can

evaluate the ground motion at the site. There are two basic steps to estimate the ground motion at the

site.

1. i Identification of the important characteristics of the ground motion (i.e. the intensity measures). The selection of the IM of interest (amplitude, frequency and/or duration-based) depends on the element at risk under consideration and on the purposes of the analysis.
2. 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 beside). In probabilistic terms, this is written as P[IM > imjR = r;M = m] = 1 􀀀 FIMjRM(imjr;m).

In general, the conditional distribution of the ground motion intensity measure, i.e. F(imjr;m) is assumed

log normal. It follow that the logarithm of the conditional intensity measure is normally distributed.

Some observations to mention. Well, over decades of development, the prediction models have become complex, consisting of many parameters. More speci\_cally, many factors related to source and site characteristics

are considered and among all the parameters involved, faulting mechanism effect of local site conditions are of paramount importance.

-----------------------------------------------------------------------------------------------------------------------------------------------

Then finally the last step: the hazard computation.

The seismic hazard curve is a function representing the annual frequency of exceeding various levels of

ground shaking (i.e. the IM) at a speci\_c site. The curve is obtained by integration of the \_rst 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.

Given Ns statistically independent sources, the aggregate hazard curve is derived by merging all the Ns

Poisson process. Then, numerical integration is required.

-----------------------------------------------------------------------------------------------------------------------------------------------

Now let’s see with Matlab how to perform a PSHA analysis of a site located in Italy and then, if someone is curious, we can compare it with results from the interactive seismic hazard map of the Italian National institute INGV, by applying just some modification to our code.

The goals are here reported, so basically to compute the hazard curve for each fault annual, at 50 years or at 475 years. What we expect, reasonably, is that the last one hazard curve is going to have an higher rate of exceedance, since, as one can imagine, it has a major return period so chance to see more seismic events. Let’s try to dramatically simplified in this way our expectations.

-----------------------------------------------------------------------------------------------------------------------------------------------

As first step we have seen that we need to characterize sources involved in hazard evaluation.

To do so, we need to access for example to the database and catalogue provided by the Italian National Institute INGV. Here we found data and material needed for the localization of the seismogenetic zones. In particular, the updated catalogue and model for the Italian GMPE is the so called ZS9.

In this seismological zonation, we find different categories based on distinct characteristics such as fault mechanism, hypocenter depth and so on.

-----------------------------------------------------------------------------------------------------------------------------------------------

Here, since we placed ourselves in Palmoli, middle of Italy, we have selected some of the seismogenetic zone useful or interesting for the hazard curve calculation. In particular, from this table, some useful information are already reported, such as coefficient b of GR, m max and m min ect.

-----------------------------------------------------------------------------------------------------------------------------------------------

Besides, in this first phase, also geometrical information about how to model the seismic source are needed: in the same catalogue are reported the coordinates that define the polygon of the area source.

These, for calculation purposes, are then converted in a more suitable format and reported in a cartesian reference system.

-----------------------------------------------------------------------------------------------------------------------------------------------

In this way, it is possible to compute the pdf distribution of site-to-source distance. Just for sake of clarity, here is reported a conventional circular zonation: the maths in this way is simplified, as the definition of the CDF corresponds to an analytically easy close form. We can therefore interpret the CDF as the ratio between area of circle with radius small r over area of circle of fixed radius, in this case 50 km.

So, since pdf is the derivative of CDF, we obtain a linear probability distribution of site-to-source distance.

-----------------------------------------------------------------------------------------------------------------------------------------------

These steps corresponds to the following sections in the codes provided. So there is a pre-processing section in the m-file inside which data from the INGV catalogue are read and prepared for further elaboration.

-----------------------------------------------------------------------------------------------------------------------------------------------

As It possible to see, the pdf of distance is discretized and plotted. Plots show the effect of how a low or bad step of integration can influence results and convergence. As always there’s a penalty to pain: increasing number steps causes the computational effort needed, so it has to be an informed decision.

-----------------------------------------------------------------------------------------------------------------------------------------------

Then, as second step, it is possible to evaluate the pdf distribution of magnitude by considering GR bounded law. Again, starting from CDF and by remembering the definition, it is possible to derive also the pdf.

-----------------------------------------------------------------------------------------------------------------------------------------------

Again, here are plotted the continuous and discretized form of the PDF and CDF for magnitude.

The codes reported are simply the formulation rewritten in matlab.

-----------------------------------------------------------------------------------------------------------------------------------------------

At this stage, we need to implement the most delicate phase of the hazard analysis: the prediction equation or the conditional probability that we have seen before. GMPE are model capable to predict the pdf of ground motion intensity as function of predictor variables such as magnitude, distance and other parameters. Generally, they are based on the form here reported.

-----------------------------------------------------------------------------------------------------------------------------------------------

Just for simplicity, here is reported the predictive equation of Cornell that is for PGA intensity measure. It is a formulation functions of magnitude and site distance with some constants value calibrated.

Just an observation: in the Italian database Sabetta Pugliese law is the one applied to evaluate the hazard map. In codes you will find the possibility to change or implement that: it requires more data that are collected all in the documentation provided by the INGV.

-----------------------------------------------------------------------------------------------------------------------------------------------

We can thereby compute the probability of exceeding any PGA level by evaluating the conditional probability applying for example the Cornell law. As we noticed, the conditional probability can be evaluated as the complement of the cumulative distribution function.

To give a visual idea of what we are computing we can give a look to the figure below: for a given magnitude and distance, we are interested in the probability of the intensity measure PGA, so the pdf here plotted for different distances, greater than a threshold: we are looking for the coloured part of the drawn pdf.

-----------------------------------------------------------------------------------------------------------------------------------------------

Here the coded example of the attenuation law of Cornell

-----------------------------------------------------------------------------------------------------------------------------------------------

So finally we need to compute the hazard by combining all the previous info.

We compute the probability of exceedance first, then the rate by multipling for the rate of exceedance of earthquakes for the selected source.

-----------------------------------------------------------------------------------------------------------------------------------------------

And finally we consider all the involved sources in the hazard computation.

-----------------------------------------------------------------------------------------------------------------------------------------------

By discretizing for the numerical solution

-----------------------------------------------------------------------------------------------------------------------------------------------

So here are reported the solution we can obtain by running the provided code in matlab.

-----------------------------------------------------------------------------------------------------------------------------------------------

Welcome back. Let’s focus now on the fragility analysis.

-----------------------------------------------------------------------------------------------------------------------------------------------

Fragility analysis represent the next step on the PBEE framework that we have seen in the first part of the lecture.

Its definition is below reported: again, fragility is defined as the conditional probability of an event, a defined limit state, given the observation of an intensity measure which describe the seismic event.

-----------------------------------------------------------------------------------------------------------------------------------------------

Fragility analysis involved both vulnerability and damage analysis, since it requires the evaluation of both probability of EDPconditioned by IM and DM conditioned by EDP.

In fact, by definition, the fragility curve is defined as the conditional probability of failure of a structure or its critical component, at given values of seismic intensity.

In practice a fragility is calculated as the conditional probability tha the damage measure D exceeds a critical threshold for a given seismic IM.

----------------------------------------------------------------------------------------------------------------------------------------------- So, here is reported an example of fragility curves for different damage limit states, defined according to some criteria based upon the engineering demand parameters, for incresing level of pga.

-----------------------------------------------------------------------------------------------------------------------------------------------

We understand therefore how intensity measure and damage state definition are crucial in the formulation of fragility functions. In greater details, for the intensity measure some properties here enlisted are required: for example efficiency, robustness, practicality syffucuebct abd effectiveness.

Regarding damage state definition instead, it should suit the specific structural problem. In other terms, it is fundamental that damage state definition are based on edp that are significative for the structural problem. Generally, the damage states are de\_ned with categorical variables as follow

\_ D0 No damage; \_ D1 Slight, minor damage; \_ D2 Moderate damage;\_ D3 Extensive Damage and C Collapse.

It must be said that it is also possible to de\_ne continues damage states, e.g. through continuous damage indexes.

-----------------------------------------------------------------------------------------------------------------------------------------------

There are generally four macro-classes of fragility functions:

i Empirical. Empirical fragility are obtained by fitting a function to observational data from past earthquake events or laboratory tests. Data are usually collected as ordered pairs of level of excitation and categorical variables of damage or collapse. Usually, these methods are based on pure statistical analysis (e.g. regression models).

Analytical fragility functions are derived by de\_ning an analytical, or numerical,

structural model and analyzing its performance under di\_erent levels of the seismic

hazard. There are two subclasses of analytical models: static, i.e. hazard considered as response spectrum and the structural analysis performi s usualy a pushover analysis; dinamic i.e. a collection of gms and simulation via non linear analysis on fe models;

Expert opinion fragilities is created by polling one or more (independently)

experts of the given structural asset. Usually, it is asked to guess or estimate the

failure probability, or a range of failure probabilities, for a given hazard level.

Hybrid fragility functions are obtained based on a combination of the di\_erent

methods.

-----------------------------------------------------------------------------------------------------------------------------------------------

Incremental dynamic analysis IDA [?] is probably the most popular method to compute fragility

functions via time-history analysis. Briey stated IDA can be summarized as follows

-----------------------------------------------------------------------------------------------------------------------------------------------

Then finally, perform the computation of the fragility according to different formulation in case of full or truncated IDA as reported here below.

In the first case, since analysis are scaled until collapse is reached, all collected data are from a system that have reached collapse. So the Fragility, by assuming for example a lognormal pdf for the random variable IM, can be evaluated as the CDF of the normal distribution with parameters of sigma and mu evaluated as below.

Different is the case of the truncated or even of the cloud analysis, in which there is an upper limit on the intensity measure: so not in all of the analysis the system has reached the damage desidered. Data are then collected in data that causes collapse and in data that do not cause collapse. So, for both the families the likelihood is evaluated: in the first case as the pdf of lognormal distribution, in the second case as the complement CDF of the lognormal distribution.

-----------------------------------------------------------------------------------------------------------------------------------------------

So the likelihood of the whole dataset is composed by the product of the two contributes.

By applying log form we obtain the formulation reported in eq. (2) by which the estimation of the parameters of the distribution is gained simply by optimization of the solution of the problem.

-----------------------------------------------------------------------------------------------------------------------------------------------

More in generally, we can compute fragility in this manner by define a generic pdf: this is an equivalent way of describe the problem by not limiting our description just to a lognormal distribution.

-----------------------------------------------------------------------------------------------------------------------------------------------

So, let’s see now an application study and some code implemented inside matlab for fragility evaluations.

-----------------------------------------------------------------------------------------------------------------------------------------------

The objective is to perform fragility analysis for this case study, later on described.

So, given a provided set of ground motions, our goal is to perform both classical and truncated IDA and determine fragility curves for both the structure ATTEL along the MRF direction and the BF direction. Moreover, the implemented model consider the possibility to adopt a bouc wen model to take into account also hyesteresis.

-----------------------------------------------------------------------------------------------------------------------------------------------

So the main steps in which also codes are organized are here summarizes:

1. A definition of the numerical model of the structure is required;
2. Input and proper IM is needed, and this as we will see is obtained by scatter plot and data exploration of the dataset of gms;
3. Definition of proper limit states and informative edp for the examinated system
4. Non linear time histories analysis to be performed, via IDA or truncated IDA or several other methods;
5. Collecting EDP – IM and computing fragilities

-----------------------------------------------------------------------------------------------------------------------------------------------

Let’s start with the first step: the FE model of ATTEL, for more information in the references you will find a link to some more material about it. It consists of a multi level steel frame structure: in particular along one direction the structural system consists of a moment resisting steel frame and has a period of 2.79 s, so very flexible. Along the orthogonal one, the structural system has concrete shear walls leading to a principal period along this direction of almost 0.46 s. Here just some reference sections are reported.

-----------------------------------------------------------------------------------------------------------------------------------------------

The structural system, designed according to EC8 seismic provisions, have been implemented firstly in a high fidelity model into OpenSees. By considering beam and column elements with linear elastic behaviour, but also taking into account mechanical non-linearities inside plastic hinges through a bouc-wen modified model. Also second order effects like P-Delta have been considered in the analysis.

-----------------------------------------------------------------------------------------------------------------------------------------------

Since the high number of simulation and analysis requests, we decided to adopt a simplified model in order to reduce the computational burden and time requests. A calibration procedure oriented to correspondence of main periods, modes of vibrating have been carried out resulting in a reduced MDOF model, reported inside the codes for matlab.

-----------------------------------------------------------------------------------------------------------------------------------------------

A special attention was dedicated to modeling the hysteretic behaviour through a bouc wen model here very briefly reported. Some documentation about it is reported in the references. Just a mention, the bouc-wen model essentially consists in a first-order non-linear differential equation that relates the input displacement to the output restoring force in a hysteretic way. By choosing a set of parameters appropriately, it is possible to accommodate the response of the model to the real hysteresis loops

-----------------------------------------------------------------------------------------------------------------------------------------------

Once the structural model is defined, we need to properly prepare the seismic input. Here we have choosen a dataset from the NGA-west 2 that includes 206 ground motions.

The main features of this dataset are here summarized: first of all, the records are all from crustal earthquakes with magnitude greater than 6 and a distance from site-to-source greater than 10 km, in order to avoid near faults events or near field scenarios, shear wave velocity greater than 600 m/s. This domain of gms is then partitioned according to the fault mechanism: that could be reverse fault or strike slip.

The whole ensamble was then investigated considering different possible intensity measures here enlisted.

-----------------------------------------------------------------------------------------------------------------------------------------------

So we checked through time histories acceleration, velocity and displaments and plots of spectra before referring to scatter plot and statistical tools to find out the optimal IM capable to relationate with EDP of our considered structure.

-----------------------------------------------------------------------------------------------------------------------------------------------

In this section of the code you see the selection of the input. Given the dataset, in our tutorial we run simulations with gms having a strike-slip fault mechanism.

-----------------------------------------------------------------------------------------------------------------------------------------------

Next step is about the definition of damage limit states as function of EDP for the analysed system. As you can see here reported there are some limit states function of the interstorey drift for both MRF and BF configuration, according to the FEMA 356 standards.

-----------------------------------------------------------------------------------------------------------------------------------------------

So, in the codes we report and set all of these information: first we have to choose the structural system onto which we would like to perform the analysis, then we have to choose if we want to consider the hysteretic behaviour, that for sake of simplicity has already been calibrated in the model and provided parameters, and finally to define the limit states before of performing our simulations.

-----------------------------------------------------------------------------------------------------------------------------------------------

Here is reported the core of the analysis: the definition of the time history to assign to the structure and the computation response.

-----------------------------------------------------------------------------------------------------------------------------------------------

Finally, after performing of the analysis we can compute the fragilities by applying the equation we have seen before and that here I have reported just for clarity.

As you can seen they are different for the untruncated case, as we have seen here we apply a sort of inference on the data since every of the pairs is recording for collapse reached; while for the truncated not all of the collected data refers to reached collapse, so we apply the formulation seen before.

-----------------------------------------------------------------------------------------------------------------------------------------------

Finally I’ve plotted here the results of fragility computation for both MRF and BF system configuration considering hysteretic behaviour. Just one quick observation before shifting to codes, for the MRF, since it is very flexible, we need severe PGA in order to see high probability of collapse; instead for the BF we reached collapse for lower values.