**Abstract**

In this paper, an innovative strategy to generate broad band strong ground motion records is presented. The methodology couples the outcomes of two different analyses: (1) the wave-forms obtained by a forward deterministic physics-based simulation (PBS) of the seismic scenario and (2) the response spectra predicted through Artificial Neural Networks (ANN), trained on a selected database. The time-histories obtained with PBS embodies a detailed underlying seismological model (i.e. source-, path- and site-effects), but it is reliable in a low-frequency (LF) range (i.e. up to 1.5-2 Hz) due to computational limitations. On the other hand, ANN predicts response spectral ordinates at short-periods (SP), once trained on a set of strong ground motion records. The ANN is constrained by PBS long-period (LP) spectral ordinates and it predicts the SP ones, since PBS values have no sense in this range. Before applying the ANN to LF synthetics, the latter are enriched at high-frequencies (HF), by merging their Fourier’s spectra with empiric/stochastic ones, generated independently. Finally, the hybrid pseudo-acceleration response spectrum is iteratively scaled to match ANN prediction at short-period. The procedure is validated for two Italian strong ground motion events, whose numerical seismic scenarios were available. A comparison between recorded and simulated intensity map at short-period is presented.

**Introduction**

An accurate estimation of the ground motion is a key issue for an adequate assessment of the seismic hazard at the site of interest. In fact, aiming at bridging earthquake information from earth sciences and the engineering practice for mitigating the earthquake hazard (Aki, 2003) and forced by the complexity of ground excitation records, strong-motion seismology firstly proposed to describe the earthquake shaking by means of some key ground motion parameters (e.g., amplitude, frequency content, and duration) and linking these latter with the damages produced to structures and facilities.

The concept of elastic response spectrum was introduced by Maurice A. Biot (Biot 1933, 1934, 1941) and it becomes a fundamental method in earthquake engineering (Housner 1941; Housner et al., 1953; Hudson, 1962). Since the early ‘40s at Caltech, George W. Housner began to publish calculations of response spectra from accelerographs, nevertheless for design purposes it turned out to be more appropriate to use a collection of different recorded ground motions in order to create the so called “design spetcrum” as already suggested by Biot in 1941:“When we possess a collection of earthquake spectrums at a given location, it is suggested that a simplified envelope should be used as a standard spectrum for the purpose of design in that region.” Newmark and Hall (1982) proposed a design spectrum which has been used extensively in research and engineering practice: the procedure starts with peak values of ground acceleration, velocity and displacement. Also for this reason, the ever since core business of strong-motion seismology has been the reliable prediction of peak time-domain amplitudes.

To this end, recorded seismograms of increasingly good quality are commonly made available after major events, allowing, among various other purposes, to calibrate ground-motion prediction equations (GMPE, or “attenuation” relationship) across the entire frequency band of engineering interest. GMPE are usually referred to as *empirical* models, since the calibration is based primarily on empirical ground motion data and provides a statistical estimate of the expected ground motion and the standard deviation by means of certain earthquake parameters (e.g.: magnitude, distance, soil condition , etc.). It is worth to mention that during the last decades, some authors proposed to use GMPEs calibrated on a combination of both empirical and numerically simulated ground motion data to overcome the limitation related to available observed data (Abrahmson&Silva 2007).

Due to simplicity and limited computational cost, the aforementioned GMPEs are vastly used to estimate the ground shaking for a given scenario or to assess the seismic hazard within a given region and furthermore they play a crucial role, overall steering the seismic design procedures, and specifically prescribing the design spectrum that should be adopted for earthquake-resistant design of structures.

Although the concept of design spectrum is undoubtedly a convenient tool, the statistical analysis of the collection of spectra derived from recorded ground motions in different earthquakes, is going inevitably to loose some of their characteristics. Thus, despite the overall effectiveness and simplicity of the GMPEs, this empirical approach has major shortcomings:

* classical GMPEs predicts peak values, whereas time-histories should be preferred;
* the available records adopted to calibrate a GMPE are not always adequate to match the needs of major potential engineering applications (e.g. there are a few records available for large earthquake in near-field conditions);
* the data-driven prediction implies that the coefficients will vary when calibration databases are updated and that only a somehow generic site condition can be represented (defined, for instance, by means of Vs30 ), bypassing crucial site-specific features (e.g.: complex geological context);
* the point-wise prediction provided cannot reproduce ground motion specific spatial variability and/or coherency;
* the shape of the predicted spectra tends to be mostly dependent on magnitude and to a minor extent is related to the distance.

Peak values predicted through GMPEs may be also obtained by a *stochastic* approach. For instance, various formulas from random vibration theory will often suffice for applications requiring only peak motions. Hanks and McGuire (reference) used such an approach in their prediction of peak acceleration. Boore (reference) extended this technique, by exploiting the statistical moments of the squared amplitude spectra.

One of the major improvement due to the stochastic approach resides in its capability of generate time-histories. The idea beneath it is rather simple: the earthquake ground motion is interpreted as a non-stationary time-varying stochastic process and the synthetic spectrum is generally expressed as frequency-domain scaling function of the source intensity. Some models exploiting time and frequency modulating functions have been proposed, as far as time-varying autoregressive moving average (ARMA) model etc (reference). It is worth to remark that those approaches are still data-driven, being empirically calibrated upon available seismic databases and albeit the boost of worldwide strong-motion data resources, certain magnitude-distance ranges or specific site conditions remain poorly covered.

An interesting approach was also proposed by Sabetta and Pugliese (reference) to study the attenuation of response spectra and to simulate artificial accelerograms as a function of magnitude, distance, and site geology (based on the Italian strong-motion data). Using multiple regressions, the authors developed empirical predictive equations for the vertical and horizontal components of response spectra corresponding, between 0.25 and 25 Hz. The great advantage of this class of empirical/stochastic methods is the limited number of parameters required, easily retrieved from recorded wave-forms or from available seismological information at the site of interest. Basic yet generic information from other sites with similar features can be used for this purpose.

However, more refined stochastic approaches require a deeper insight on the underlying seismological model of the radiated spectra. The *stochastic method* proposed by Boore, for instance, filters a suite of windowed, stochastic time series so that the amplitude spectra are equal, on the average, to the specified spectra. Originally developed for a point-source, the stochastic method has been extended to finite-fault simulations, i.e. by considering each sub-fault as a point-source and convolving all the contributions. Besides peak acceleration, the model fits all the essential aspects of high-frequency (HF) ground motions (i.e. above 0.1 Hz). It has been tested on hundreds of recordings from earthquakes with seismic moments spanning more than 12 orders of magnitude and in diverse tectonic environments. However, the main parameters that affect ground motions (source, path, and site) are still cast into simple functional forms. This inevitably leads to major approximations, especially in complex near-fault conditions or when non-linear site-effects are observed.

In recent year a great interest has been given to deterministic numerical simulations of seismic scenarios, i.e. for the so called *forward Physics-Based Simulations* (PBS).

The ever increasing computational power (e.g. massively parallel supercomputers) made a 3D numerical simulations of the source-to-site seismic wave propagation at local or regional scale relatively fast and efficient. For this reason, PBS is rapidly becoming the leading and most reliable tool to construct ground shaking scenarios for future earthquakes. The method encompasses (1) the detailed description of the rupture process (kinematic or dynamic), along with (2) the detailed 3D description of the geological profile, topography and eventually (3) non-linear site effects.

Although very powerful, PBS has some major drawbacks: (1) the computational burden increases as long as the mesh becomes finer (for instance, to describe complex geological configurations); a restriction to low-frequency (LF) analysis (nowadays reliable engineering applications hardly overcome the limit of 2-3 Hz ); (2) a deep knowledge of the area under investigation is required, thus leading to the assessment of a large set of parameters (e.g. the geological properties of soil layer and crustal bedrock, the description of the fault rupture) in order to depict a reliable seismic scenario.

Physics-based modelling already proved in the past decades to be well suited for global (references) and regional scale simulation (references), nevertheless only few numerical codes worldwide (references) challenged this physically-based technique to study within a single numerical model the dynamic problem from the source of earthquake up to the structures. Even with large computational resources available, this kind of studies remain extremely challenging due to remarkable difference in terms of spatial and temporal scale required to accomplish the simultaneous analysis of this domain. Even neglecting the aforementioned computational challenge, the high-frequency band in PBS remains essentially an opened issue, due to the fact that at the moment is not possible resolving the source rupture process above 1 Hz and furthermore there is no chance to characterize the mechanical properties for large areas with the kind of accuracy that should be required in order to physically follow the process of propagation of the waves.

Having said that, realistic broad-band (BB) signals (i.e. between 0 and 25 Hz) can be obtained by the enrichment of LF records based on PBS. Several successful attempts have been made to *hybridize* LF record with HF predictions, obtained separately with empirical/stochastic methods (references). The two wave-forms are convolved together by *sewing* the LF and HF Fourier’s spectra, high-pass and low-pass filtered respectively. No phase processing is performed on the two records. Unfortunately, hybrid wave-forms barely reflect the motion spatial coherence and near-field features due to the mentioned limitations of empirical/stochastic approaches. In this sense, final response spectra at long-period (LP) are intrinsically incompatible with short-period (SP) ones.

In this paper we propose an alternative strategy for spectral ordinate estimation at short period. The described approach is based on the combination of physics-based simulation with Artificial Neural Networks (ANN). An ANN is a data-processing algorithm (or an actual hardware) designed upon biological neural networks. In other words, it can be regarded as a statistical-computational framework set up to artificially reproduce the predictive capability of interconnected logic units (i.e. neurons). A neural network is constructed by organizing sets of interconnected neurons in layers to estimate or approximate non-linear functions that usually depends on a large number of inputs. Due to their adaptive interconnection, the neurons *learn* from examples, and exhibit some structural capability for generalization. Moreover, neural networks normally have great potential for parallelism, since the computations of the components are independent of each other. An ANN has the following basic features: it contains sets of adaptive weights, i.e. numerical parameters that are tuned by a learning algorithm and it is capable of approximating non-linear functions of their inputs. The adaptive weights can be thought of as connection strengths between neurons, which are activated during the *training* phase (upon a set of data, compatible with the expected outcome) and exploited during the *prediction* phase.  
In this context, ANNs are trained upon exemplary strong ground motion database, taking LP spectral ordinates as input data (once a corner period has been arbitrarily chosen). Applying the trained ANN upon the synthetic time-histories provided by deterministic numerical simulations (PBS or hybridized), one gets two response spectra, exactly coincident at long-period, but diverging at short-periods. Therefore, by scaling the synthetic response spectrum upon ANN prediction at short periods, one gets broad-band synthetics whose response spectrum is predicted by ANN constrained by deterministic analysis of the earthquake scenario.  
In Section [sec:syn2ann\_ann\_training], the ANN training procedure is described in detail. The design of the training workflow is presented at first: neural networks are constructed by testing their regression performance on validation sets. In the following, two crucial aspects are discussed: the sensitivity to the site-class of the training dataset and the predictive performance with respect to the vertical component of motion. Section [sec:syn2ann\_recipe] outlines the recipe to produce broad-band strong ground motions from 3D physics-based numerical simulations, following the workflow depicted in Figure [fig:syn2ann\_flow\_chart]). Some applications of the methodology are also presented in the following. Specifically, two Italian strong ground motion earthquakes are considered, namely: (1) the 2009 M*W*6.3 L’Aquila earthquake, Central Italy and (2) M*W*6.0 2012 May 29 earthquake in the Po Plain, Northern Italy. Recording stations nearby the epicentres were considered in this study, since they experienced an intense ground shaking in near-field conditions.

Aki, K. (2003). A perspective on the history of strong motion seismology, Phys. Earth. and Planet. Int. 137,5.11.

Abrahamson & Silva NGA Ground Motion Relations for the Geometric Mean Horizontal Component of Peak and Spectral Ground Motion Parameters. Norman A. Abrahamson

PEER Report 200x/xx, Pacific Earthquake Engineering Research Center, College of Engineering, University of California, Berkeley, July 9, 2007, Draft Version 2

Biot, M.A. (1933). Theory of elastic systems vibrating under transient impulse with an application to earthquake-proof buildings. Proceedings, National Academy of Sciences , 19, 262–268.

Biot, M.A. (1934). Theory of vibration of buildings during earthquake. Zeitschrift für angewandte Mathematik und Mechanik , 14, 213–223.

Biot, M.A. (1941). A mechanical analyzer for the prediction of earthquake stresses. Bulletin of the Seismological Society of America , 31, 151–171.

Housner, G.W. (1941). Calculating the response of an oscillator to arbitrary ground motion. Bulletin of the Seismological Society of America , 31, 143–149.

Housner, G. W. (1952) Spectrum Intensities of Strong-Motion Earthquakes. In: Proceedings of the Symposium on Earthquake and Blast Effects on Structures : Los Angeles, California, June 1952. Earthquake Engineering

Hudson, D.E. (1962). Some problems in the application of spectrum techniques to strong-motion earthquake analysis. Bulletin of Seismological Society Of America, 52, 417–430.

Newmark, N.M. and Hall, W.J. (1982).Earthquake Spectra and Design. Monograph, Earthquake Engineering Research Institute, Berkeley, California.

Some links:

N. M. Newmark and R. Riddell, 'Inelastic spectra for seismic design', Proc. 7th worldconf. eurthquuke eng. Istanbul, Turkey 4, 129- eng. Istanbul, Turkey (1980).

<http://www.iitk.ac.in/nicee/wcee/article/7_vol4_129.pdf>

<http://peer.berkeley.edu/course_modules/eqrd/index.htm?c227top.htm&227cont.htm&ElResp/elresp5.htm>

Riddel:

<http://www.iitk.ac.in/nicee/wcee/article/11_2126.PDF>

<http://www.ce.berkeley.edu/~mahin/CE227web/Bozorgni-CampbellCh_BerteroBozorgnia.pdf>

Biot:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1085941/?page=6>

[http://www.ce.berkeley.edu/~mahin/CE227web/Sect\_7\_LinearRespSpec\_09-1.pdf](http://www.ce.berkeley.edu/%7Emahin/CE227web/Sect_7_LinearRespSpec_09-1.pdf)

<https://www.researchgate.net/post/What_is_difference_between_time_history_analysis_and_response_spectrum_analysis>

<http://www.curee.org/image_gallery/calendar/essays/1998-CUREE_excerpt.pdf>

<http://peer.berkeley.edu/pdf/AbraSil_2007_v2.pdf>

<http://e-collection.library.ethz.ch/eserv/eth:14476/eth-14476-01.pdf>

<http://www.uni-potsdam.de/nadi/docs/kuehn_diss.pdf>

<https://de.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/55887/versions/1/screenshot.png>