

Generator of genetic seismic signals

Silvia García* and Leonardo Alcántara

Received: September 27, 2017; accepted: March 22, 2019; published on line: July 01, 2019

Resumen

En este trabajo se presenta un método original para el ensamblado de acelerogramas compatibles con un espectro objetivo preestablecido. El procedimiento se basa en un generador de señales sísmicas, el cual toma en consideración las características del evento sísmico y las condiciones del subsuelo para hacer una representación realista de los movimientos del terreno. Se utilizan operadores genéticos para transformar en forma iterativa los acelerogramas, imitando los procesos de apareamiento, selección natural y mutación. El generador de señales sísmicas genéticas, modifica los componentes de aceleración de los registros considerados buenos padres para producir individuos exitosos o nuevas señales óptimas que mejor se ajustan a las condiciones objetivo. El procedimiento es inmediato y consistente con resultados notables en registros que coinciden con una amplia variedad de espectros objetivo, con una desviación mínima y conservando las cualidades geotécnicas y sismológicas inherentes de los padres (acelerogramas registrados).

Palabras clave: algoritmos genéticos, cómputo evolutivo, generación de señales sísmicas, sísmos sintéticos, espectro de respuesta, simulación sísmica.

Abstract

In this paper, an original approach for assembling time series of accelerations compatible with a prescribed target spectrum is presented. This generator of seismic signals takes into consideration earthquake and soil characteristics for a realistic depiction of ground motions. The proposed methodology uses genetic operators to transform the time series iteratively. Mimicking mating, natural selection, and mutation, the generator of genetic seismic signals, modifies the accelerations components of records considered good parents for producing successful individuals or new optimal signals that best fit the target conditions. The procedure is immediate and consistent with remarkable results in records that match a broad variety of target spectra with minimal deviation while conserving the geotechnical and seismological inherent qualities of the parents (recorded accelerograms)..

Key words: genetic algorithms, evolutionary computing, seismic signals generation, synthetic earthquakes, target spectrum, seismic simulation.

S. García*
Coordinación de Geotecnia
Instituto de Ingeniería
Universidad Nacional Autónoma de México
Ciudad Universitaria, 04510
Coyoacán, CDMX, Mexico
*Corresponding author: sgab@pumas.iingen.unam.mx

L. Alcántara
Coordinación de Geotecnia
Instituto de Ingeniería
Universidad Nacional Autónoma de México
Ciudad Universitaria, 04510
Coyoacán CDMX, Mexico

Introduction

Nonlinear time-history analysis is becoming an everyday demand in seismic analysis and design of structures. To specify time-series input excitations to structural models it is necessary to generate signals (artificial accelerograms) having compatible characteristics with modern evaluation guidelines. Since traditionally the seismic hazard at a site for design purposes has been represented by design spectra, virtually all seismic design codes and guidelines require scaling of selected ground motion time histories so that they match or exceed the controlling design spectrum within a period range of interest (ICC, 2000; ASCE, 2000; NEHRP 2011).

The techniques of scaling time histories can be classified as i) the frequency-domain methods, where the frequency content of the recorded ground motions are manipulated (Gasparini and Vanmarcke, 1976; Silva and Lee, 1987; Bolt and Gregor, 1993; Department of the Army, 2000; Carballo and Cornell, 2000) and ii) the time-domain methods, which limit themselves to control the amplitude of recorded ground motions (Kircher, 1993; Naeim and Kelly, 1999). Regardless of the method domain, the processes of selecting the "initial" ground motions and their scaling to match the design spectrum are separate and distinct.

The procedure presented in this paper involves a comprehensive search for time histories based on seismological and geotechnical characteristics and a genetic adjustment of the selected series for matching a target spectrum. The genetic routine proposed here considers the variability in the characteristics of the recorded strong-motions under similar geotechnical and seismic specifications and produces shaking time series using an ensemble of accelerograms rather than just one or two "typical" records for matching target spectra.

In structural and earthquake engineering, during the past decade, genetic algorithms have been used in design optimization of nonlinear structures (Pezeshk *et al.*, 1999 and Pezeshk *et al.*, 2000), active structural control (Alimoradi, 2001), and performance-based design (Matouš *et al.*, 2000; Foley and Schinler, 2001; Foley *et al.*, 2003; Shaw *et al.*, 2004). Therefore, using genetic algorithms to scale earthquake ground motions for design is but a natural continuation of such applications and parallels the attractive use of neural networks to achieve the same task (Ghaboussi and Lin, 1998; Kim and Ghaboussi, 1999).

Genetic Algorithms

In order to make this article self-contained, some of the design considerations involved in implementing a GA to solve the particular problem are discussed briefly. To the interested reader, deeper theoretical explanations by Mitchell (1996) and Michalewicz (1996) are recommended. The application of GA in engineering contexts is described in Rani and Moreira (2009).

In the computer science field of artificial intelligence, genetic algorithms GA is a searching heuristic tool that mimics the process of natural selection (Mitchell, 1996). This heuristic (also sometimes called a metaheuristic) is routinely used to produce useful solutions to optimization problems. Genetic algorithms (Holland, 1975; Goldberg, 1989) belong to the larger class of evolutionary algorithms EA, which generate solutions to optimization problems using techniques inspired by natural evolution, such as *inheritance, mutation, selection, and crossover*.

Algorithm basics. This concise explanation tails the flow chart in Figure 1. The first step is defining an objective function with inputs and outputs and the binary GA encodes the value of each input parameter (e.g. q, r, s, t) as a binary number (Figure 2). If the variable of the parameter space of an optimization problem is continuous, a real coded GA is indicated. The parameter values are then placed side-by-side in an array known as a *chromosome*. A *population* is a matrix with each row representing a *chromosome*.

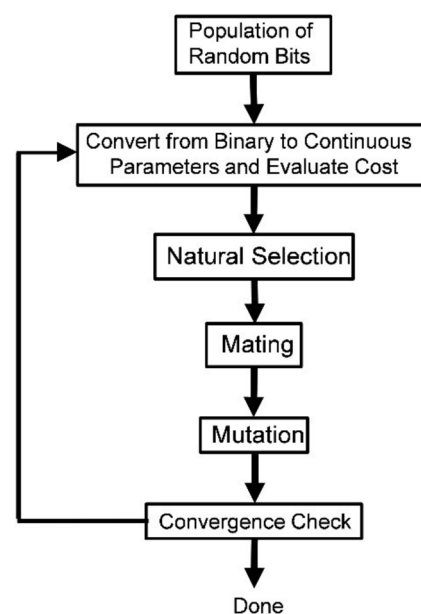


Figure 1. Flow chart of Binary Genetic Algorithm.

A population consisting of random ones and zeros will be used as an example for describing the algorithm steps (Figures 1 and 2). These random binary digits translate into guesses of values of the input parameters. Next, the binary *chromosomes* are converted to continuous values, which are evaluated by the objective function. *Mating* takes place between selected *chromosomes*. Mates are randomly selected with a probability of selection greater for those chromosomes yielding desirable output from the objective function (tournament or roulette wheel selection). *Offspring* (new chromosomes) produced from mating inherit binary codes from both parents (Figure 3a). A simple *crossover* scheme randomly picks a crossover point in the *chromosome*.

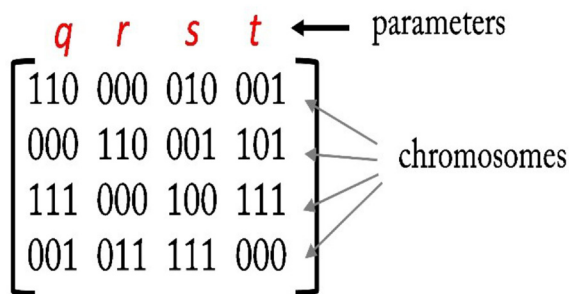


Figure 2. Initial population of binary coded parameters.

An *offspring* results by keeping the binary strings to the left of the crossover point for each parent and swapping the binary strings to the right of the crossover point. *Crossover* mimics sexual recombination in nature, where two parent solutions are chosen and parts of their subtree are swapped and because each function exhibits the property 'closure' (each tree member is able to process all possible argument values), every crossover operation should result in the formation of a legal structure. *Mutation* (Figure 3b) causes random changes in an individual before it is introduced into the subsequent population. Unlike crossover, mutation is asexual and thus only operates on one individual; during mutation all functions and all terminals are removed beneath an arbitrarily determined node and a new branch is randomly created, or a single node is swapped for another. Stronger individuals will have greater chance to survive across epochs and to reproduce than weaker individuals, which will tend to perish (Figure 3c).

The objective function outputs associated with the new population are calculated and the process repeated. The algorithm stops

after finding an acceptable solution or after completing a set number of iterations. Selecting the best population size, mating scheme and mutation rate is still an area of controversy. Since the GA is a random search, a certain population size and mutation rate can give considerably different answers for different independent runs (Haupt and Haupt, 1998 and 2000; Haupt, 2003; Ursem, 2003).

Some advantages of genetic over conventional numerical optimization algorithms are i) optimizing with continuous or discrete parameters, ii) do not require derivative information, iii) simultaneously search from a wide sampling of the objective function surface, iv) deal with a large number of parameters, v) optimize parameters with extremely complex objective function surfaces, vi) provide a list of semi-optimum parameters, not just a single solution, and vii) works with numerically generated data, experimental data, or analytical functions. These advantages outweigh the GAs' lack of rigorous convergence proofs. For thorough discussion on this subject, the interested reader is referred to Goldberg (1989), Michalewicz (1992) and Beasley *et al.* (1993).

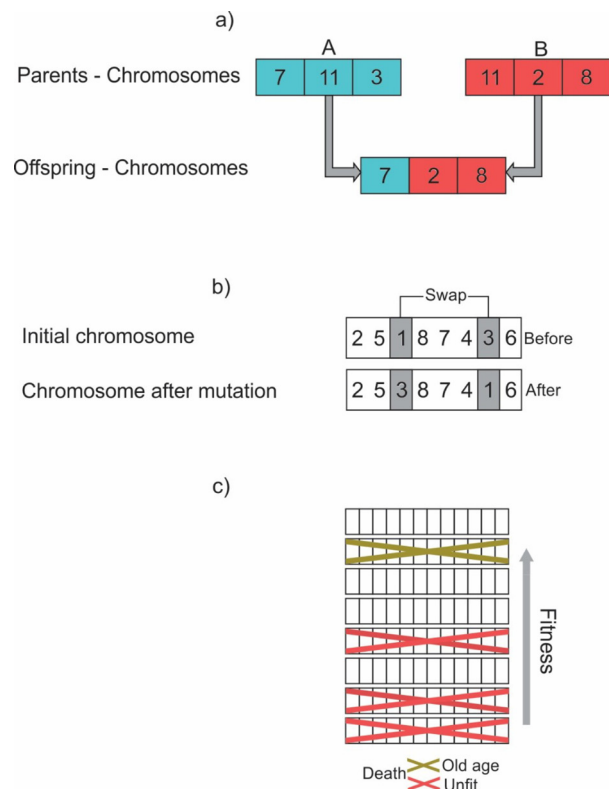


Figure 3. Genetic algorithms operations. a) Crossover, b) Mutation and c) Death.

Genetic accelerograms: the underlying idea

A certain number of inputs (set of recorded accelerograms), say x_1, x_2, \dots, x_n , belonging to the input space X is selected. Using GA terminology, each accelerogram is called an *organism* or *chromosome* and the set of *chromosomes* is designated as a *colony* or *population*. To each *chromosome* x_i a fitness value $f(x_i)$ is assigned.

The objective of using $f(x_i)$ when generating genetic accelerograms is to find the best combination of time series that minimizes the difference between a given target spectrum and the one obtained from the genetically generated accelerograms. The deviation from the target is measured by the mean square of error between the genetic spectrum (estimated from the genetic accelerograms) and the target spectrum (Figure 4). The problem is formulated as the minimization of the error function, Z , between the averaged scaled spectra and the target spectrum in a range of T_i to T_f :

$$Z = \min \left\{ \frac{1}{T_f - T_i} \sum_{i=1}^f ([SA_{gi}(T)] - [SA_i(T)])^2 \right\}$$

in which

T = the vibration period

$SA_{gi}(T)$ = spectral acceleration of genetic number at period

$SA_i(T)$ = spectral acceleration of target record number at period

T_i = initial period to consider

T_f = final period to consider

This formulation does not guarantee that the final solution would not fall below the target in the period range under consideration; instead, it would merely attempt to minimize the deviation of the solution from the target.

The algorithm will tend to keep accelerograms (*organisms*) closer to the optimal in the set of inputs (the colony) and to discard those that under-perform, analogous to *natural selection*. The crucial step in this routine is the *reproduction* or *breeding* that occurs once per epoch. The genetic information of the two accelerograms participating in reproduction are literally merged together to form a new *chromosome* (a *child* accelerogram generated by *crossover*). This heuristic allows the possibility to combine the best of both individuals to yield a better one (*evolution*). During each epoch, a given fraction of the *organisms* is allowed to *mutate*, providing

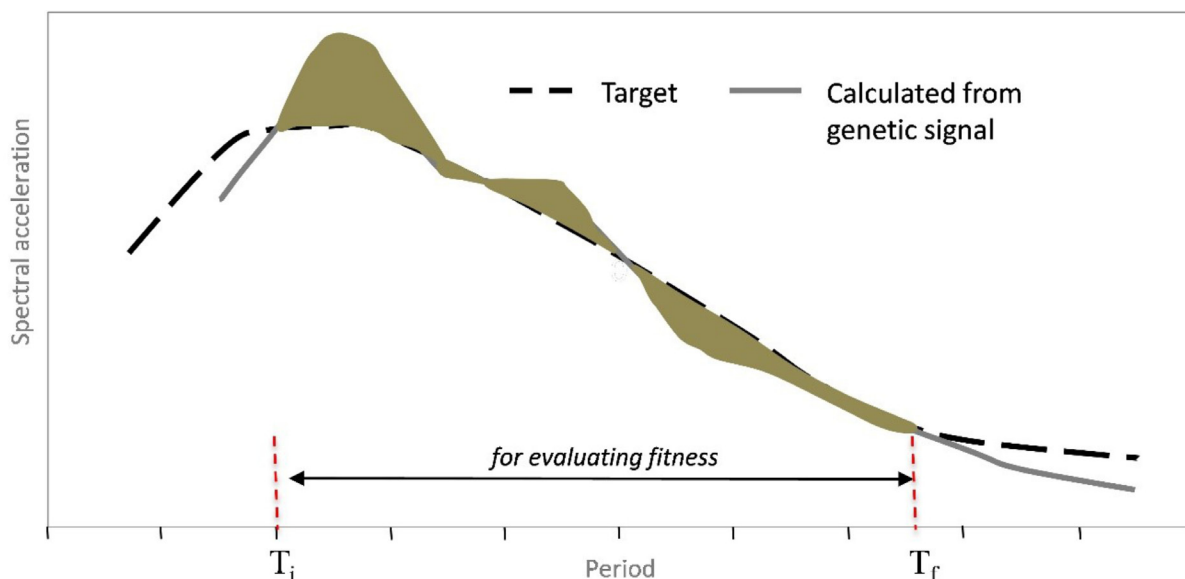


Figure 4. Fitness value: spectra representation.

randomness for spanning the whole input space and generating individuals with random genes. Worst individuals in the *population* have a high probability of dying at the end of each epoch, old chromosomes can also be eliminated and the highest performer is immune from old-age death (Figure 3c).

In this investigation, the backbone genetic algorithm routines from the GaNetXL code placed in public domain by Savić *et al.* (2011), were adapted and modified. An overall flow chart of the program operation is presented in Figure 5. For starting, additional data are required: i) the time series Δt , ii) the number of components of time series, iii) the ordinates, i.e. spectral acceleration, of the target acceleration design spectrum, iv) the matching period, v) the maximum values of differences, and vi) the set of GA parameters. The GA parameters are population size, number of generations, and crossover and mutation ratios.

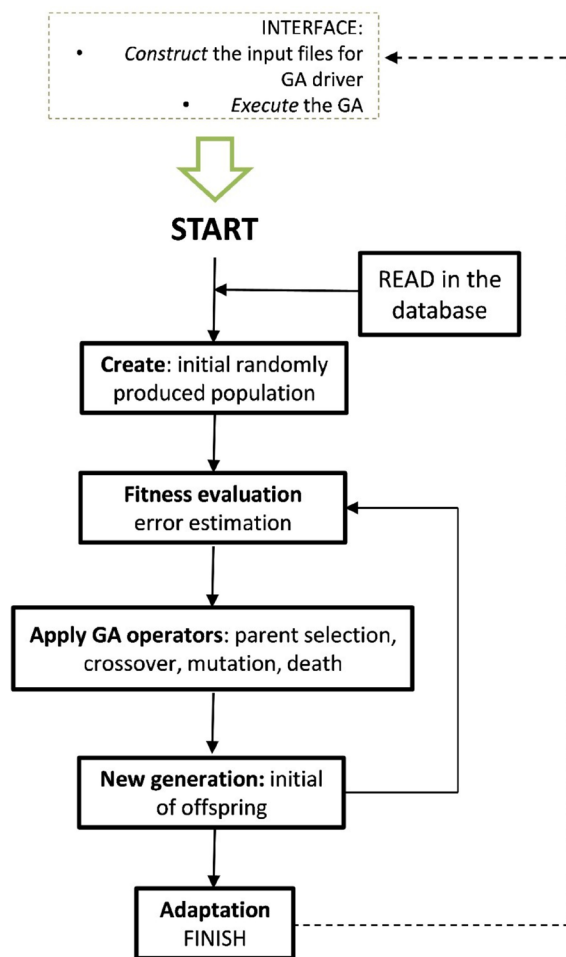


Figure 5. The program flow chart.

One of the advantages of the proposed model is that the original *chromosomes* can be selected according to geotechnical and seismic characteristics. The entire *population* is contained in the Mexican strong motion database (*Base Mexicana de datos de Sismos Fuertes* in original language) which includes more than 13 000 ground motion records (SMIS, 2000). Obviously, this population is suitable for studying seismogenesis at the Mexican subduction zone, but for different applications, any appropriately categorized set of records could be used.

The user can define from the vast universe of parents, which conditions are more relevant for a particular analysis. The system meets the request from i) one of the predefined seismic environments (Ordaz and Reyes, 1999) based on two dynamic maps of México (Figure 6), ii) magnitude M_w , iii) focal depth FD and iv) soil type in the recording station ST . Soils are classified as "Type A" for soft materials with high plastic index, high compressibility, high water content, and low to very-low shear wave velocities, and as "Type B" for deposits made up of stiff materials with high strength, high to very-high shear wave velocities and low to very-low compressibility potential. If the user does not have a priori seismic or geotechnical preferences, the genetic generator selects the initial population randomly.

Application examples

In the following, several examples are used to demonstrate the proposed method. First, three cases are presented in order to generate signals which match a given target spectrum without a priori preferences. These instances also illustrate the stability of the genetic algorithm in adapting itself to peculiar spectrum shapes. The second description demonstrates the kind of signals that are developed from preselected earthquake environment and soil types, condition especially important for detailed seismic hazard analyses. Examples are presented for two different types of soil.

Earthquake signals generated from random population in order to match prescribed spectrum shapes

The target spectra for this example are shown in Figure 7a (dashed line). The period range of 0.10 to 3.0 seconds was assumed for matching the targets. A genetic search of a 300-individual population (designated randomly) over 500 generations with a crossover ratio of 60.0% and a mutation probability of 2.0% was utilized. Figure 7b shows the fitness curve as a function

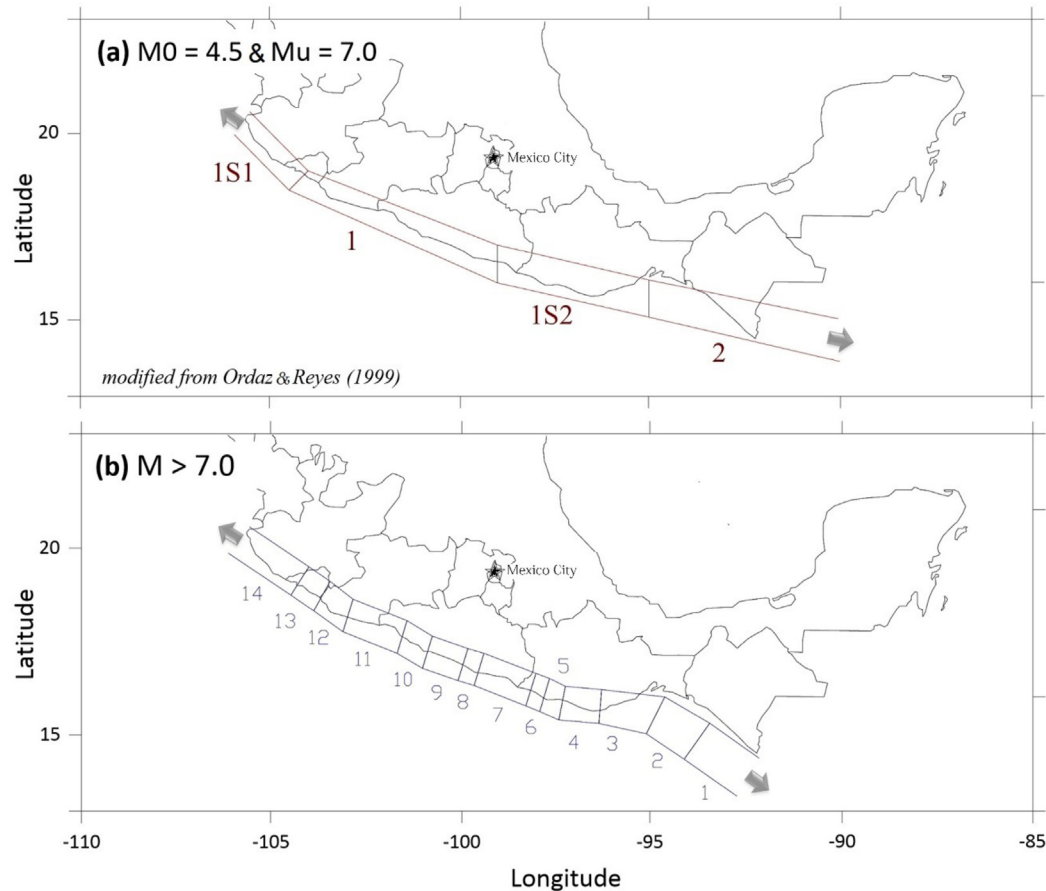


Figure 6. Predefined seismic environments for Mexican subduction.

of successive generations. After completing the iterations, the signals recovered following this procedure are samples that illustrate the application of the genetic methodology to create accelerations time series that match a smooth, of the type included in building codes or steep spectrum shapes.

The genetic algorithm generates three records representing the best match to the target spectra and have the highest fitness levels (Figure 7c). As can be seen in Figure 7d, the differences between the spectra from the genetically generated records and the target spectra in the range of 0.1 to 3.0 seconds are minor. The genetic signals resemble the target curve with remarkable accuracy.

Earthquake signals generated from population constrained by prescribed source and soil-type

In this case, seismic and geotechnical conditions are specified. These conditions are Type B soils (deposits of very dense sand, gravel, or very stiff clay-ground), seismicity

from Zone 1 (see selected zone in Figure 6a), records with magnitudes between $6 < M < 7$ and *undefined* focal depth (no restrictions). The initial population was set at 100 individuals. The target peak ground acceleration PGA has been set equal to 0.028g. After 2780 generations, the stop criteria is achieved and the spectra from the *offspring* can be evaluated. The response spectra of the simulated artificial earthquakes are compared with the target response spectrum in Figure 8. The good match shows the accuracy of the proposed procedure and the suitability of the generated artificial accelerograms for precise design purposes.

The next example has the following restrictions: Type A (soft clays), seismicity from Zone 2, $5 < M < 6.5$, $10 < FD < 35$ km and initial population of 50 individuals. For this requirement, and after 3100 generations, the genetically generated spectral accelerations attain slightly higher values than the target spectrum at some frequencies. The genetically generated spectral accelerations, however, lies very near the target around the natural frequency (Figure 9).

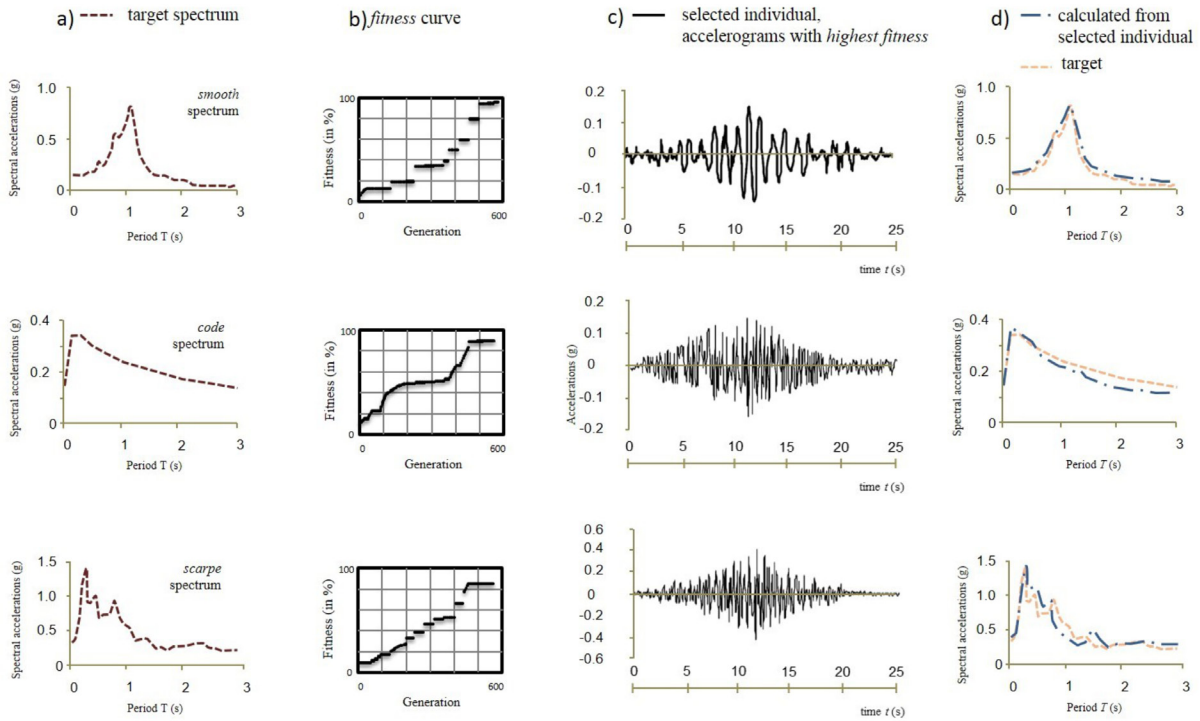


Figure 7. Some examples of genetic earthquake signals matching prescribed spectra.

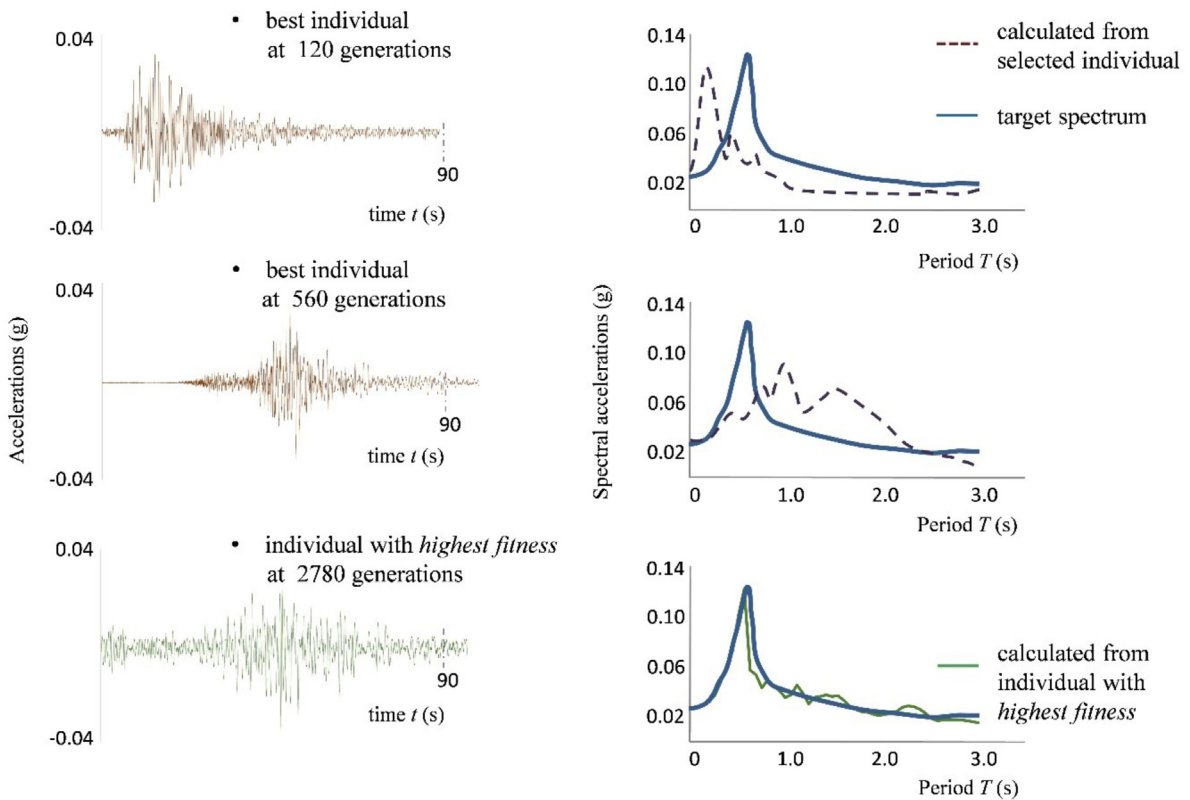


Figure 8. Evolution of offsprings matching prescribed seismic and geotechnical conditions for Type B soils.

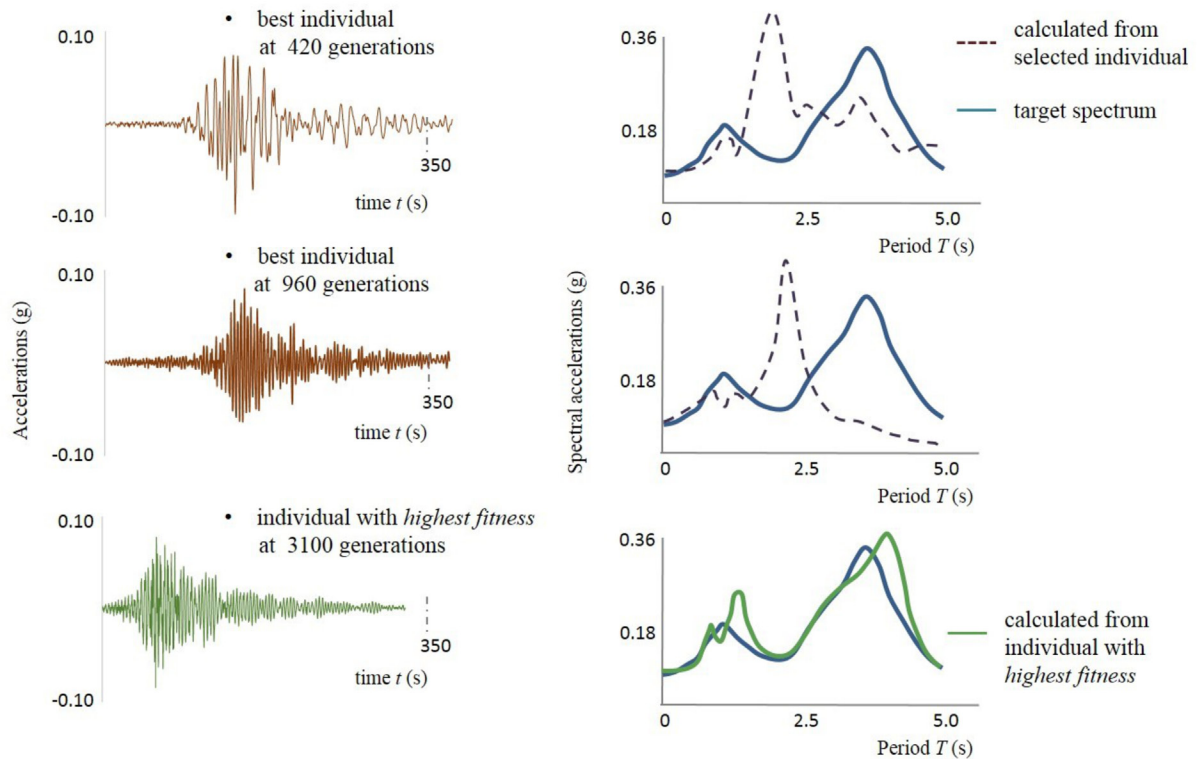


Figure 9. Evolution of offsprings matching prescribed seismic and geotechnical conditions for Type A soils.

Conclusions

A new method for transformation of earthquake ground motions for matching a specific response spectrum was presented. Given a target spectrum, a processes inspired in Darwin's theory about evolution is used to generate artificial time histories from recordings. The procedure is fast and reliable and produces records matching the target spectrum with minimal deviation. The method uses a search engine for selecting a first generation of individuals, and a genetic routine for modifying the initial population through the processes that mimic *mating*, *natural selection*, and *mutation*. The process continues until an optimum individual (best *fitness*) is obtained.

References

- Alimoradi, A., 2001. Performance Study of a GA-Based Active/Hybrid Control System Under Near Source Strong Ground Motion, *2001: A Structural Engineering Odyssey, Proceedings of the 2001 Structures Congress and Exposition*, Washington, D.C., ASCE.
- American Society of Civil Engineers (ASCE), 2000. Prestandard and Commentary for the Seismic Rehabilitation of Buildings, prepared for the SAC Joint Venture, published by the Federal Emergency Management Agency, FEMA-356, Washington, D.C.
- Beasley, D., Bull, D.R., Martin, R.R. 1993. An overview of genetic algorithms: Part 1, fundamentals *University Comput.*, 15 (2), pp. 58–69.
- Bolt, B. A., and Gregor, N. J., 1993. Synthesized strong ground motions for the seismic condition assessment of the eastern portion of the San Francisco Bay Bridge, Report UCB/EERC- 93/12, University of California, Earthquake Engineering Research Center, Berkeley, CA.
- Carballo, J. E., and Cornell, C. A., 2000. Probabilistic Seismic Demand Analysis: Spectrum Matching and Design, Department of Civil and Environmental Engineering, Stanford University, Report No. RMS-41.
- Department of the Army, 2000. Dynamic Analysis of Concrete Hydraulic Structures, Circular No. EC 1110-2-6051, Washington, D.C.
- Foley, C. M., and Schinler D., 2001. Optimized Design of Partially and fully-Restrained

- Steel Frames Using Distributed Plasticity, *2001: A Structural Engineering Odyssey: Proceedings of the 2001 Structures Congress and Exposition*, Washington, D.C., ASCE.
- Foley, C. M., Pezeshk, S., and Alimoradi, A., 2003. State of the Art in Performance-Based Design Optimization, *2003: Proceedings of the ASCE Structures Congress*, Seattle, Washington, ASCE.
- Gasparini, D., and Vanmarcke, E. H., 1976. SIMQKE: A Program for Artificial Motion Generation, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- Ghaboussi, J., and Lin, C.J., 1998. New Method of Generating Spectrum Compatible Accelerograms using Neural Networks, *Earthquake Engineering & Structural Dynamics*, 27(4): 377-396.
- Goldberg, D. E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley.
- Haupt, R. L., and Haupt, S. E., 1998. *Practical Genetic Algorithms*, John Wiley and Sons, New York.
- Haupt, R. L., and Haupt, S. E., 2000. Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors, *Appl. Computat. Electromagnet. Soc. J.*, 15(2): 1034-1037.
- Haupt, S. E., 2003. Genetic Algorithms in Geophysical Fluid Dynamics, AMS Conference on Artificial Intelligence, P1.7
- Holland, J. H. *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI: University of Michigan Press 1975.
- International Code Council (ICC), 2000. *International Building Code*, Falls Church, Virginia.
- Kim, Y., and Ghaboussi, J., 1999. A New Method of Reduced-Order Feedback Control Using Genetic Algorithms, *Earthquake Engineering & Structural Dynamics*, 28(3): 235-254.
- Kircher, C., 1993. Personal communication with Farzad Naeim and Marshall Lew. Levy, S., 1992. *Artificial Life*, Vintage Books, New York.
- Matouš, K., Lepš, M., Zeman J., Šejnoha M., 2000. Applying genetic algorithms to selected topics commonly encountered in engineering practice, *Comput. Methods Appl. Mech. Engrg.* 190: 1629-1650.
- Michalewicz, Z., 1992. *Genetic Algorithms + Data Structures=Evolution Programs*, 1st ed., Springer, Berlin.
- Michalewicz, Z., 1996. *Genetic Algorithms + Data Structures=Evolution Programs*, 3rd ed., Springer, Berlin.
- Mitchell, M., 1996. *An Introduction to Genetic Algorithms*, MIT Press, London.
- Naeim, F., and Kelly, J. M., 1999. *Design of Seismic Isolated Structures—From Theory to Practice*, John Wiley and Sons, New York.
- National Earthquake Hazards Reduction Program (NEHRP), 2011. *Earthquake-Resistant Design Concepts: An Introduction to the NEHRP Recommended Seismic Provisions for New Buildings and Other Structures*, FEMA P-749.
- Ordaz, M and C Reyes (1999), "Earthquake hazard in Mexico City: Observations versus computations", *Bulletin of the Seismic Society of America*, 89(5): 1379-1383.
- Pezeshk, S., Camp, C. V., and Chen, D., 1999. Genetic Algorithm for Design of Nonlinear Framed Structures, *Proceedings of the 1999 Structures Congress*, New Orleans, Louisiana, ASCE.
- Pezeshk, S., Camp, C. V., and Chen, D., 2000. Design of Framed Structures by Genetic Optimization, *ASCE Journal of Structural Engineering*, 126(3), Paper No. 19107, 382-388.
- Rani, D., Moreira, M. M., 2009. Simulation-Optimization Modeling: A survey and potential application in reservoir systems operation. *Water Resources Management*, 24(6): 1107-1138.
- Savić, D. A., Bicić, J., & Morley, M. S. 2011. A DSS Generator for Multiobjective Optimization of Spreadsheet-Based Models. *Environmental Modelling and Software*, 26(5): 551-561. <http://dx.doi.org/10.1016/j.envsoft.2010.11.004>.
- Shaw, D., Miles, J., Grey, A., Genetic programming within Civil Engineering organization of the adaptive computing in design and manufacture 2004 Conference, Engineers House, Clifton, Bristol, U.K.

Silva, W., and Lee, K. 1987. State-of-the-art for assessing earthquake hazards in the United States, Report 24, ES RASCAL Code for Synthesizing Earthquake Ground Motions, Miscellaneous Paper S-73-1, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Sociedad Mexicana de Ingeniería Sísmica (SMIS), 2000. Mexican strong motion database CD-Rom, vol 2.

Ursem, R. K., 2003. Models for evolutionary algorithms and their applications in system identification and control optimization, BRICS Dissertation Series DS-03-6, ISSN 1396-7002.