Earthquake Risk Models with Genetic Algorithm Hybridization

Blind Author 1 University 1
Address 1
Address 1
email@address.com

Blind Author 1 University 1
Address 1
Address 1
email@address.com

Blind Author 1 University 1
Address 1
Address 1
email@address.com

ABSTRACT

The GAModel is a method which aims to generate forecasts by using only Evolutionary Computation (EC). However, the GAModel was limited by the very high number of parameters, and the subsequent large search space. This project main goal is to refine the GAModel ideas, objective to have a greater ability to predict the future behavior of groups of earthquakes events by overcoming this limitations.

This document sumamarizes the GAModel and proposes three new method based on the GAModel. The first method is called the ReducedGAModel. By this, we expecte to reduce the impact of the amount of earthquakes parameters and the size of the search space generating a forecast model faster. For that we used a different genome representation than the representation used in the GAModel. In the GAModel, the genome is a real valued genome, with every gene coresponds to a specific bin in the model. In the ReducedGAModel, the genome is a list of earthquake events locatation. Each element of the genome has a correspondent bin in the model.

We also wanted to do a hybridization, a association of EC and geophysical knowledge, of the GAModel and the ReducedGAModel with empirical laws, such as the modified Omori-Utsu formula. These new methods names are Emp-GAModel and Emp-ReducedGAModel. This hybridization is performed in two phases. The first one, is to obtain forecast models obtained by the GAModel and the ReducedGAModel. After that, in the second phase, these models are all refined by the same group of geophysical formulas. We expect that we will be able to generate not only faster model but more accurated ones, because we will reduce the search space and increase its learning rate with the formulas.

The models generated by these four method were evaluated and compared based on the predictabily experiments framework proposed by the "Collaboratory for the Study of Earthquake Predictability" (CSEP), an international effort to standardize the study and testing of earthquake forecasting models. The experiments were designed to compare

1-year earthquake rate forecasts for four regions in Japan in using the data from the Japan Meteorological Agency (JMA) earthquake catalog.

1. INTRODUCTION

Earthquakes may cause soil rupture or movement, tsunamis and more. They may cause great losses and that can be explicit by some examples, such as the earthquakes in Tohoku (2011) and Nepal(2015). To be able to minimize the consequences of these events, we look to create forecast earthquake occurences models. Hence the characteristics that most influence the earthquakes events may vary both in time and place, these methods should be to adapt their behavior to be able to forecast earthquakes events which reflects well the reality.

This project aims to obtain a better method, based in improvements to the GAModel [1], a statistical method of analysis of earthquakes risk using the Genetic Algorithm technique (GA). Two ideias are proposed for this. The first one, is to change the candidate solution representation. By that, we objective to make the GAModel more specialized, focusing only on areas on which earthquakes happened already in a near past. This will lead to a faster convergence, once the amount of parameters is smaller and consequently, the search space gets smaller.

Formulated on this idea, we propose the ReducedGAModel. Its genome only has information of areas that already had occurences in the past. This helps the method to converge gets faster, by minimizing the number of parameters the method has to deal with.

The other ideia is based on the assumption that earthquakes cluster in both space and time, and the we want associate the Genetic Algorithm technique (GA) with a some empirical laws, such as the modified Omori law. First, the background intensity (the independent earthquakes or mainshocks), which is a function of the space, is forecasted using the GA. Then, we use some empirical laws to obtain the dependent earthquakes (aftershocks) for a specific time interval.

The Emp-GAModel is the method proposed that incorpores some geophysical knowledge. It is a hybridization of the models generated by the GAModel with the these empirical laws, see Section 3.

Finally, there is the Emp-ReducedGAModel. This method is a combination of the two ideias. Therefore, it also performs a hybridization of models with the group of empirical law. Though, for this method, the models are generated by the ReducedGAModel method and not by the GAModel.

The forecast models produced by those methods and the ones produced by the GAModel were all analyzed using likelihood tests, namely the L-test, the N-test and the R-test, as suggested by Regional Earthquake Likelihood Model (RELM) [21].

For developing the methods and to be able to compare them we used the earthquake catalog from the Japanese Meteorological Agency (JMA), using event data from 2005 to 2010.

This paper is organized as: in Section ??reviews applications of Evolutionary Computation in the context of seismology research. The next Section,Section 3, we give a details of each of the forecast proposed covering the Collaboratory for the Study of Earthquake Predictability (CSEP) framework and the empirical laws. In Section 4, we give the description of the tests proposed in [22]. After that, in 5, we define the target areas used for the experiment and the data from the JMA; we clarify the design followed during the experiments and how ew compared the forecast models derived from our methods. Finally, we show the results and conclude this work in 6 and 7.

2. EVOLUTIONARY COMPUTATION FOR EARTHQUAKE RISK ANALYSIS

labelestadoArte Reports of the application of Evolutionary Computation and related methods for the generation of earthquake forecasts are rather sparse.

The usage of Evolutionary Computation in the field of earthquake risk models is somewhat sporadic. Zhang and Wang [27] used Genetic Algorithms to fine tune an Artificial Neural Network (ANN), and use this system to produce a forecast. Zhou and Zu [28] also proposed a combination of ANN and EC, but their system only forecasts the magnitude parameter of earthquakes. Sadat, in the paper [19], used a very simliar ideia to predict the magnitude of the earthquakes in North Iran, but in thisc ase he used ANN and GA.

Some sismological models were developed aiming to estimate parameter values by using Evolutionarry Computation.

For example, Evolutionary Computation was used to estimate the peak ground acceleration of seismically active areas [12, 3, 10, 11]. Ramos [18] used Genetic Algorithms to decide the location of sensing stations and Saeidian [20] made a comparison in performance between the GA and Bees Algorithm to decide which of those techniques would performe better when chosing the location of sensing stations. Nicknam et al. [15] and Kennett and Sambridge [9] used evolutionary computation to determine the Fault Model parameters of a earthquake.

2.1 What is Evolutionary Computation And What are Genetic Algorithms

Evolutionary Computation (EC) is concerned with algorithms based on the Darwinian principles of natural selection. It may find, by try trial and error and based on a great amount of data, better solutions for problems that human beings may not find it easy to solve [13]. That could also be done without any domain knowledge about the problem to be controlled [14].

The main goal of a Genetic Algorithm (GA), a EC technique, is to find approximated solutions in problems of search

and optimization. Based on Koza [13], GA are mechanism of search based on natural selection and genetic. They explore historical data to find optimum search points with some performance increment [6].

A GA uses those mechanisms to generate solutions to optimization and search problems. The first step is to create an initial population of possible solutions, where each solution is called an individual. Those individuals have its fitness value estimated and those with greater fitness value are then chosen to reproduction. After some evaluations, we expect to find an optimum solution.

Frequently, the initial population is randomly generated once it is common to ignore the main aspects that influence the algorithm performance. In other words, hence it is common to lack any domain knowledge, the random population is a good way to search for optimum solutions.

3. THE FORECAST MODELS USING GENETIC ALGORITHM

All forecast models proposed in this paper are based in the Collaboratory for the Study of Earthquake Predictability (CSEP) framework.

Each individual has its own representation of the framework based on different perceptions of what are the best aspects of the framework.

The population is trained on earthquake event data for a training period, which is anterior to the target test period. After completing the evalution limit, the best individual is chosen to be the final forecast.

icio

3.1 1-year Models

In the CSEP framework, a forecast model uses one common format in the literature, which is a gridded rate forecast [26]. For this format a geographical region, during a start date and an end date is divided in sections, the bins. The forecast will estimate the number (and sometimes the magnitude) of earthquakes that happens in this target region, during the target time interval, considered to be of one year for this study [1].

Large and independent earthquakes, also known as mainshocks, are followed by a wave of others earthquakes, the aftershocks [23]. Hence there is no physical measurement to identify mainshocks and its aftershocks [23], we divided the forecast models in two groups: the ones that only forecasts mainshocks and those that forecast both mainshocks and aftershocks.

3.2 Genome Representation

The GAModel each individual represents an entire forecast model [1]. For the ReducedGAModel and each individual represents a subarea of the forecast model. This subarea is related to earthquakes past events locations only, so the genome size is usually smaller than the one used in the GAModel and the Emp-GAModel.

The Emp-ReducedGAModel and the Emp-GAModel differs only from the ReducedGAmodel and from the GAmodel, respectively, by the use of equations after the forecast is provided. This means that the theirs genome representation are the same as the GAModel and the ReducedGAModel,correspondingly.

For all methods, the genome is a real valued array X, where each element corresponds to one bin in the desired

model (the number of bins n is defined by the problem). Each element $x_i \in X$ takes a value from $[0, 1)$. In the initial population, these values are sampled from a uniform distribution. For more details of the genome represatation, please refer to [1].

3.3 Fitness Function

For the fitness function we used the log-likelihood value. The fittest individual among all the others, is preserved in the next generation, to make the solution of one generation as good as the its last generation. The bins, a gene (a element) of the genome representation, b_n, define the set β and n is the size of the set β:

$$\beta := b_1, b_2, ..., b_n, n = |\beta| \tag{1}$$

The probability values of the model j, expressed by the symbol Λ, is made of expectations λ_i^j by bin b_i. The vector is define as:

$$\Lambda^j = \left(\lambda_1^j, \lambda_2^j, \ldots, \lambda_i^j\right); \lambda_i^j := \lambda_i^j(b_i), b_i \in \beta \tag{2}$$

The vector of earthquake quantity expectations is defined as: earthquake by time. The Ω vector Ăľ composed by observations ω_i per bin b_i, as the Λ vector:

$$\Omega = \left(\omega_1, \omega_2, \ldots, \omega_i\right); \omega_i = \omega_i(b_i), b_i \in \beta \tag{3}$$

The calculation of the log-likelihood value for the ω_i observation with a given expectation λ is defined as:

$$L(\omega_i|\lambda_i^j) = -\lambda_i^j + \omega_i \log \lambda_i^j - \log \omega_i! \tag{4}$$

The joint probability is the product of the likelihood of each bin, so the logarithm $L(\Omega|\Lambda^j)$ is the sum of for $L(\omega_i|\lambda_i^j)$ every bin b_i:

$$
\begin{aligned}
L^j = L(\Omega|\Lambda^j) &= \sum_{i=1}^{n} L(\omega_i|\lambda_i^j) \\
&= \sum_{i=1}^{n} -\lambda_i^j + \omega_i \log \lambda_i^j - \log \omega_i!
\end{aligned}
\tag{5}
$$

The fitness function is a coded version of the equation 5. It uses the probabilities of the bins of each individual of model for the λ values.

3.4 Evolutionary Operators

Both the GAModel and the Emp-GAModel uses a combination of operators made available by the Distributed Evolutionary Algorithms in Python (DEAP) [4]. We use the One Point Crossover for the crossover operator, for the mutation operator, the Polynomial Bounded Mutation and for selection, we use Elitism and Tournament selection, with size 3.

For the ReducedGAModel and the Emp-ReducedGAModel, the only different operator is the mutation fuction. We use a simple mutation operator which samples entirely two new valus. The first, is a new real value from $[0,1)$ and the seconde one, a new integer value from $[0,X]$, where X is the maximum size of the genome.

3.5 Mainshock Models

The GAModel is purely based on the framework suggested by the CSEP. In it, one forecast is defined as a place for a specfic time and is divided in bins. Each bin represents a geographical interval. The whole target area of study is covered by these bins and represents the $\mu(x, y)$, the background intensity [29]. In the GAModel, each possible solution is represented as an entire forecast model.

In this context the GAModel is considered as one method to generate space-rate-time forecasts. It also could be described as:

$$\lambda(t, x, y, M|\Upsilon_t) = \mu(x, y) \tag{6}$$

where you can denote the number of earthquakes forecast in a bin as $\lambda(t, x, y)$ [26] given that Υ_t is the earthquake observation data up to time t.

The ReducedGAModel, which represents the first ideia (see 1), is a method described as the GAModel, but each possible solution represents a fraction of the forecast where we expect to find especific risk areas. The GAModel defines a expected number of earthquakes for every single bin in the target region. That could lead to exhaustive and, sometimes worthless, searches. That is caused by the number of bins in the forecast and also because some in some bins there are no earthquake occurances in the observation data. To minimize this, the ReducedGAModel only define expected number of earthquakes in bins that already had some occurance in the past, giving some sort of guideline to the GA.

To make it clear, we use the same example as the one used in [1].The "Kanto" region, one of the four areas used in this study, is divided into 2025 bins (a grid of 45x45 squares). Each bin has an area of approximately $25km^2$. The GAModel then calculates an expected number of earthquakes for every bin on a determinated time interval, so the GA searches for good values in 2025 bins.

The ReducedGAModel will first obtain the position of past occurances and will calculate some expected number of earthquakes for only the bins related to those positions. For example, if there are 10 bins with occurances in Kanto in the last year, it will make the GA search good values for only those 10 bins, leaving the other 2015 bins with the value zero, representing zero occurances. It is important to highlight that in the worst case, it will make the same amount of searches as the GAModel.

3.6 Mainshock+Aftershock Models

Hence earthquakes cluster in space and inspired by the space-time epidemic-type aftershock sequence (ETAS), the Emp-GAModel, represents the second ideia (see 1) and is described as:

$$\lambda(t, x, y, M|\Upsilon_t) = \mu(x, y)J(M) \tag{7}$$

$$\lambda(t, x, y|\Upsilon_t) = \mu(x, y) + \sum_{t_i \in t} K(M_i)g(t - t_i)F \tag{8}$$

The Emp-GAModel uses $\mu(x, y)$ as defined for the GAModel, so it calculates an expected number of earthquakes for every bin in the target region. The Omori law, $g(t)$, which is considered one empirical formula of great success [29] [24] [17], is a power law that relates the earthquake occurance and its magnitude with the decay of aftershocks activity with time.

For this approach we used the probabilty density function (pdf) form of the modifed Omori law [29]:

$$g(t) = \frac{(p-1)}{c(1 + \frac{t}{c})^{(} - p)}$$

(9)

In the paper [24], Utsu says that most p and c values, for various earthquake data sets fall in the range between 0.9 and 1.4, and between 0.003 and 0.3 days, respectively. These values were based on the Davidon-Fletcher-Powell optimization procedure and used in ETAS [24]. For the experiments done for this paper, we choose the values of 1.1 for p and 0.003 for c, arbitrarily.

For $K(M_i)$, the total amount of triggered events, we count aftershocks within a calculated aftershock area, A, using the formula, where M_c is the magnitude threshold:

$$K(M_i) = A \, exp([\alpha(M - M_c)])$$

(10)

From [16], states alpha as the inverse of the magnitude of an event, or $magnitide^{-1}$. To obtain A, the following equation from [25], was used:

$$A = e^{(1.02M-4)}$$

(11)

and lastly, the $J(M)$ is a simulation of the event magnitude by Gutenberg-Richter's Law, using 1 as the value of β [7]:

$$J(M) = \beta e^{-\beta(M - M_c)}, M \geq M_c$$

(12)

At last, the Emp-ReducedGAModel is a mix between the two ideias, so it is represented as the Emp-GAModel but its candidates take the same form as the ones in the ReducedGAModel.

4. TESTS FOR EVALUATING MODELS

In the paper *Earthquake Likelihood Model Testing* [21] is proposed some statistical tests that are used in this study to compare and evaluate the forecast models, developed by the The Regional Earthquake Likelihood Models (RELM). These tests are based on the log-likelihood score that compares the probability of the model with the observed events.

To evelute the data-consistency of the forecast models we used the N-Test, the Number Test, and the L-Test, or Likelihood Test. These tests fall are significance tests. Therfore, assuming a given forecast model as the null hypothesys, the distribuition of an observable test is simulated. If the observed test statistic falls into the upper or lower tail of this distribuition, the forecast is rejected [23].

To be able to compare the model that passed the N-Test and the L-test, the R-Test, hypotheses Comparison Test is used. It calculates the relative performance of a model, by comparing the Log-likelihood values between two forecast models.

4.1 Likelihood Test or L-Test

The L(ikelihood) Test considers that the likelihood value of the model is consistent with the value obtain with the simulations. The value is calculated by the formula, where $\widehat{L}_k$ is the value of the Log-likelihood of the model j, in the *bin i* and $\widetilde{L}$ is the value of the Log-likelihood of the simulation j in the *bin q*:

$$\gamma_q^j = \frac{\left| \left\{ \widehat{L}_k^j | \widehat{L}_k^j \leq \widetilde{L}_q^j, \widehat{L}_k^j \in \widehat{L}^j, \widetilde{L}_q^j \in \widetilde{L}^j \right\} \right|}{|\widehat{L}^j|}$$

(13)

The analisys of the results can be splited into 3 categories, as follows:

1. Case 1: γ^j is a low value, or in other words, the Log-likelihood of the model is lower then most of the Log-likelihood of the simulations. In this case, the model is rejected.

2. Case 2: γ^j falls near the half of the values obtained from the simluations and is consistent with the data.

3. Case 3: γ^j is high. This means that the Log-likelihood of the data da is higher that the Log-likelihood of the model and no conclusion can be made what so ever.

It is important to highlight that no model should be reject in case 3, if based only on the L-Test. In this case the consistency can or cannot be real, therefore these model should be tested by the N-Test so that further conclusions can be done.

4.2 Number test or N-Test

The N(umber)-Test also analises the consistency of the model, but here, it compares the number os observations with the number of events of the simulations. This test is necessary to supply the underpredicting problem, which may pass unnotecied by the L-Test.

This mesures is estimated by the fraction of the total number of observations by the total number of observations of the model.

As the L-test, if the number of events falls near the half of the values of the distruition, then the model is consistent with the observation, nor estimating too much events nor few of them.

4.3 Hypotheses Comparison Test or R-Test

The Hypotheses Comparison, or the R(atio)-Test, compares two forecast models against themselves. The log-likelihood is calculted for both models and then the difference between, the observed likelihood ratio, this value indicates which one of the model better fits the observations.

The likelihood ratio is calculated for each simulated catalog. If the fraction of simulated likelihood ratios less than the observed likelihood ratio is very small, the model is reject. To make this test impartial, not given an advantage to any model, this procedure is applied symmetrically [23].

4.4 Evaluation

The evaluation process is made as follow: First, the data-consitency is tested by the L-Test and the R-test. If the model passes these tests, meaning that it was not rejected by them, it is compared with other forecast models, which also were not reject, with the R-Test. The model that best fits the R-Test is then chose as the best model [21].

5. EXPERIMENTS

To analyze the performance of the forecasts generated by the GAModel, the ReducedGAModel, the Emp-GAModel

and finally the Emp-ReducedGAModel, we used the evalution method proposed by [21] and described in section 4.4.

Objecting a better understand of the patterns that most influence the earthquakes events and also to be able to determine the qualities of those forecasts, the data of the JMA catalog was divided into four groups. Each group constituent only of earthquakes that happened in a specific time interval for a given area of Japan. The experimental data will be described in details subsequently.

5.1 Experimental Data

The data used in these experiments comes from the Japan Meteorological Agency's (JMA) catalog. It is a list of earthquakes events which took place in Japan from 2000 to 2013. Each event is characterized by some typical earhtquake information such as magnitude, latitude, longitude, and depth.

For the experiments we consider events with magnitude above 5.0 which happened in four specific areas of Japan, during the year of 2010. Those areas (Kanto, Tohoku, Kansai and East Japan) represent different earthquake attributes and could lead to more information about the power of the forecasts and/or its pitfalls. Kanto, Touhoku and Kansai contain mainly inland earthquakes, which are considered to follow more stable patterns. East Japan includes also many off-shore earthquakes. For more information about the four regions as well as a map which locates them in Japan, please refer to [1].

5.2 Experimental Design

To compare the performance of the forecast models, we execute a simulation experiment on ??? scenarios. Each scneario is defined by a region for a 1-year interval, starting in Jan/01 and endding in Dec/31. To build one scenario, one region is chosen from the set of regions (Kanto, Tohoku, East Japan and Kansai) for a determinated year. To train the methods, we use 5 years of prior data, for example, to train the 2010 1-year GAModel Kanto region, the data used is from the events which occurred in the years of 2006, 2007, 2008 and 2009 in Kanto.

To The results of the forecasts io, we used two approaches. The first one is based on the evaluation method proposed by [22] and used in [23]. As described in 4.4, the forecasts were first analysed using the L- and N- test and ?????? was chosen by the R-test as the recommended one.

The second approach is to using the Log-likelihood values obtained by each of the forecast for each scenario. The log likelihood indicates how close the forecast is to the test data, in terms of location and quantity of earthquakes [1].

All forecast models here presented are stochastic methods. Therefore, to be able to compare test the statistical significance of the results, we run each forecast model 10 times. That means that all results showed are the mean of these 10 executions.

5.3 Parameter Tuning

Parameter tunning is a commonly practiced approach which aims to find optimum parameters values before the run of a algorithm. Later this algorithm is executed with these values, which remain fixed during the run [5].

To find a good set of parameters for the all the forecast models algorithms, the pySMAC, a Python wrapper for the hyperparameter optimization tool Sequential Model- based Algorithm Configuration (SMAC) [8], was used.

SMAC is an algorithm configurator takes as input an algorithm executable, a formal description of the parameters for the algorithm, and a set of training problem instances. It searches the space of possible parameter values by generating a candidate configuration and evaluating the configuration on the set of the training instances [2].

The configuration which best fits each one of the forecast models algorithm is then chosen to be used as the set of values for the its algorithms run. For all algorithms present here, the default control parameter values (which are based on the values applied in the GAModel [1]), the ranges for the parameters (arbitrarily chosen), as well as the values found by SMAC, are shown in Table **??** for GAModel and Emp-GAmodel and in Table 5.3 for ReducedGAModel and Emp-ReducedGAModel.

EXPLICAR A QUESTAO DA EXPLICACAO DOS PARAMETROS $N_GEN * POP = CONST$

Table 1: The default control parameter values and the best set of parameters found by tunning the algorithm with SMAC for the ReducedGAModel and the Emp-ReducedGAModel

	Parameter	Range	Default	Tuned
	Parameter	Range	Default	Tuned
Kanto	Crossover Chance	[0,1]	0.9	0.9
	Mutation Chance	[0,1]	0.1	0.1
	Population Size	N/A	500	
	Generation Number	[50,250]	100	100
	Parameter	Range	Default	Tuned
Kansai	Crossover Chance	[0,1]	0.9	0.9
	Mutation Chance	[0,1]	0.1	0.1
	Population Size	N/A	500	
	Generation Number	[50,250]	100	100
	Parameter	Range	Default	Tuned
Tohoku	Crossover Chance	[0,1]	0.9	0.9
	Mutation Chance	[0,1]	0.1	0.1
	Population Size	N/A	500	
	Generation Number	[50,250]	100	100
	Parameter	Range	Default	Tuned
East Japan	Crossover Chance	[0,1]	0.9	0.8826
	Mutation Chance	[0,1]	0.1	0.1724
	Population Size	N/A	500	
	Generation Number	[50,250]	100	106

6. RESULTS

7. CONCLUSIONS

Acknowledgments

Bogdan, Zechar, Zhuang
???

8. REFERENCES

[1] C. Aranha, Y. C. Lavinas, M. Ladeira, and B. Enescu. Is it possible to generate good earthquake risk models using genetic algorithms? In *Proceedings of the International Conference on Evolutionary Computation Theory and Applications*, pages 49–58, 2014.

Table 2: The default control parameter values and the best set of parameters found by tunning the algorithm with SMAC for the ReducedGAModel and the Emp-ReducedGAModel

	Parameter	Range	Default	Tuned
Kanto	Crossover Chance	[0,1]	0.9	0.9
	Mutation Chance	[0,1]	0.1	0.1
	Population Size	N/A	500	
	Generation Number	[50,250]	100	100
	Parameter	Range	Default	Tuned
Kansai	Crossover Chance	[0,1]	0.9	0.9
	Mutation Chance	[0,1]	0.1	0.1
	Population Size	N/A	500	
	Generation Number	[50,250]	100	100
	Parameter	Range	Default	Tuned
Tohoku	Crossover Chance	[0,1]	0.9	0.9
	Mutation Chance	[0,1]	0.1	0.1
	Population Size	N/A	500	
	Generation Number	[50,250]	100	100
	Parameter	Range	Default	Tuned
East Japan	Crossover Chance	[0,1]	0.9	0.8826
	Mutation Chance	[0,1]	0.1	0.1724
	Population Size	N/A	500	
	Generation Number	[50,250]	100	106

[2] C. Aranha, R. Tanabe, R. Chassagne, and A. Fukunaga. Optimization of oil reservoir models using tuned evolutionary algorithms and adaptive differential evolution. In *Evolutionary Computation (CEC), 2015 IEEE Congress on*, pages 877–884. IEEE, 2015.

[3] A. F. Cabalar and A. Cevik. Genetic programming-based attenuation relationship: An application of recent earthquakes in turkey. *Computers and Geosciences*, 35:1884–1896, October 2009.

[4] F.-M. De Rainville, F.-A. Fortin, M.-A. Gardner, M. Parizeau, and C. Gagné. Deap: A python framework for evolutionary algorithms. In *Proceedings of the Fourteenth International Conference on Genetic and Evolutionary Computation Conference Companion*, GECCO Companion '12, pages 85–92, New York, NY, USA, 2012. ACM.

[5] A. E. Eiben and S. K. Smit. Parameter tuning for configuring and analyzing evolutionary algorithms. *Swarm and Evolutionary Computation*, 1(1):19–31, 2011.

[6] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989.

[7] A. Helmstetter and D. Sornette. Predictability in the epidemic-type aftershock sequence model of interacting triggered seismicity. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 108(B10), 2003.

[8] F. Hutter, H. H. Hoos, and K. Leyton-Brown. Sequential model-based optimization for general algorithm configuration (extended version). Technical report, Technical Report TR-2010-10, University of British Columbia, Computer Science, 2010.

[9] B. L. N. Kennet and M. S. Sambridge. Earthquake location âĂŤ genetic algorithms for teleseisms. *Physics of the Earth and Planetary Interiors*, 75(1–3):103–110, December 1992.

[10] T. Kerh, D. Gunaratnam, and Y. Chan. Neural computing with genetic algorithm in evaluating potentially hazardous metropolitan areas result from earthquake. *Neural Comput. Appl.*, 19(4):521–529, June 2010.

[11] T. Kerh, Y.-H. Su, and A. Mosallam. Incorporating global search capability of a genetic algorithm into neural computing to model seismic records and soil test data. *Neural Computing and Applications*, pages 1–12, 2015.

[12] E. Kermani, Y. Jafarian, and M. H. Baziar. New predictive models for the v_{max}/a_{max} ratio of strong ground motions using genetic programming. *International Journal of Civil Engineering*, 7(4):236–247, December 2009.

[13] J. R. Koza, M. A. Keane, and M. J. Streeter. Genetic programming's human-competitive results. *IEEE Intelligent Systems*, pages 25–31, 2003.

[14] D. Michie, D. J. Spiegelhalter, and C. Taylor. Machine learning, neural and statistical classification, 1994.

[15] A. Nicknam, R. Abbasnia, Y. Eslamian, M. Bozorgnasab, and E. A. Mosabbeb. Source parameters estimation of 2003 bam earthquake mw 6.5 using empirical green's function method, based on an evolutionary approach. *J. Earth Syst. Sci.*, 119(3):383–396, June 2010.

[16] Y. Ogata and J. Zhuang. Space–time etas models and an improved extension. *Tectonophysics*, 413(1):13–23, 2006.

[17] F. Omori. On the after-shocks of earthquakes. 1895.

[18] J. I. E. Ramos and R. A. Vázques. Locating seismic-sense stations through genetic algorithms. In *Proceedings of the GECCO'11*, pages 941–948, Dublin, Ireland, July 2011. ACM.

[19] N. Sadat, S. Zakeri, and S. Pashazadeh. Application of neural network based on genetic algorithm in predicting magnitude of earthquake in north tabriz fault (nw iran). *Current Science (00113891)*, 109(9), 2015.

[20] B. Saeidian, M. S. Mesgari, and M. Ghodousi. Evaluation and comparison of genetic algorithm and bees algorithm for location-allocation of earthquake relief centers. *International Journal of Disaster Risk Reduction*, 2016.

[21] D. Schorlemmer, M. Gerstenberger, S. Wiemer, D. Jackson, and D. A. Rhoades. Earthquake likelihood model testing. *Seismological Research Letters*, 78(1):17–29, 2007.

[22] D. Schorlemmer, M. Gerstenberger, S. Wiemer, D. Jackson, and D. A. Rhoades. Earthquake likelihood model testing. *Seismological Research Letters*, 78(1):17–29, 2007.

[23] D. Schorlemmer, J. D. Zechar, M. J. Werner, E. H. Field, D. D. Jackson, T. H. Jordan, and R. W. Group. First results of the regional earthquake likelihood models experiment. *Pure and Applied Geophysics*, 167(8-9):859–876, 2010.

[24] T. Utsu and Y. Ogata. The centenary of the omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*, 43(1):1–33, 1995.

[25] Y. Yamanaka and K. Shimazaki. Scaling relationship between the number of aftershocks and the size of the main shock. *Journal of Physics of the Earth*, 38(4):305–324, 1990.

[26] J. D. Zechar. Evaluating earthquake predictions and earthquake forecasts: A guide for students and new researchers. *Community Online Resource for Statistical Seismicity Analysis*, pages 1–26, 2010.

[27] Q. Zhang and C. Wang. Using genetic algorithms to optimize artificial neural network: a case study on earthquake prediction. In *Second International Conference on Genetic and Evolutionary Computing*, pages 128–131. IEEE, 2012.

[28] F. Zhou and X. Zhu. Earthquake prediction based on lm-bp neural network. In X. Liu and Y. Ye, editors, *Proceedings of the 9th International Symposium on Linear Drives for Industry Applications, Volume 1*, volume 270 of *Lecture Notes in Electrical Engineering*, pages 13–20. Springer Berlin Heidelberg, 2014.

[29] J. Zhuang, Y. Ogata, and D. Vere-Jones. Analyzing earthquake clustering features by using stochastic reconstruction. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 109(B5), 2004.