Improving the Generation of Earthquake Risk Models Using Evolutionary Algorithms Tempered by Domain Knowledge

Yuri Lavinas
University of Brasilia
Computer Science Department
yclavinas@gmail.com

Xiucai Ye
University of Tsukuba
Graduate School of SIE
yexiucai@mma.cs.tsukuba.ac.jp

Marcelo Ladeira
University of Brasilia
Computer Science Department
mladeira@unb.br

Claus Aranha
University of Tsukuba
Graduate School of SIE
caranha@cs.tsukuba.ac.jp

***Abstract*—Earthquake Risk Models describe the risk of occurrence of seismic events on a given area based on information such as past earthquakes in nearby regions and the seismic properties of the area under study. These models can be used to help to better understand earthquakes, their patterns and their mechanisms.**

In previous work, we showed that Genetic Algorithms (GA) could generate risk models with the same degree of precision as the Relative Intensity (RI) method, which is considered a benchmark for this problem. However, a few shortcomings were also defined in that approach: (1) The representation of the model in the Genetic Algorithm was too sparse, (2) Domain knowledge was not used to create the model, and (3) The relationship between foreshocks and aftershocks were not taken into account.

In this work, we try to address these three concerns. We propose a new representation of a seismic risk model to be used as the genome of the Genetic Algorithm. We introduces a hybrid model that incorporates seismic theories about earthquake distribution (such as the Omori-Utsu formula). And we use clustering to filter the earthquake catalog in order to remove earthquakes that are likely to be aftershocks before generating the risk model.

We examine each of these changes through simulations using the catalog of Japanese earthquakes between 2000 and 2010. According to our results, clustring the earthquake catalog produces better models, while the proposed changes to representation did not show such a clear effect. These results allow us to draw recommendations for future developments.

1. Introduction

Earthquakes can cause great damage to human society through soil rupture, movement, tsunami, etc. Some recent earthquakes that highlight this destructive potential are the great East Japan Earthquake of 2011 (depicted in figure1), and the April 2015 earthquake in Nepal. One important tool for the enactment of policies that minimize the consequences of these events are earthquake occurrence models (also called risk models). These models can be used to identify patterns in the seismic mechanisms that generate earthquakes, and are important to increase our understanding of these events.

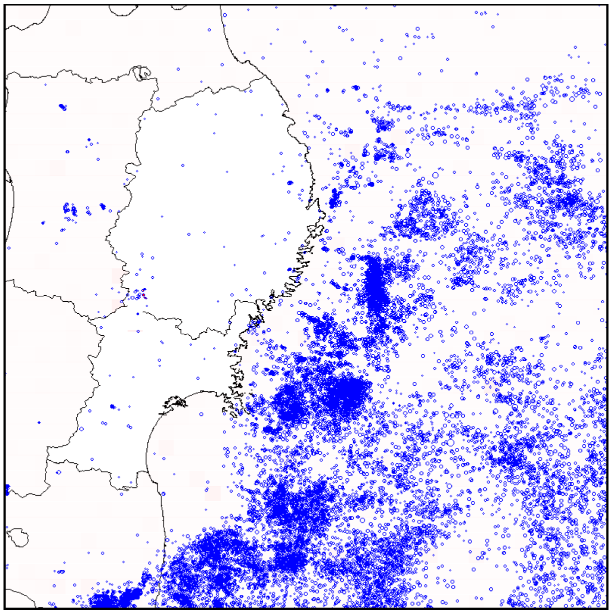

Figure 1. Seismic Activity in Eastern Japan in 2011. Each blue dot represents one earthquake

In our previous work [1], we proposed a way to generate earthquake risk models using a standard Genetic Algorithm (here called the GAModel). The GA model was shown to be competitive with the Relative Intensity (RI) model, while not using any a-priori information about the distribution of earthquake occurrences (See section 2 for a summary of the GAModel, along with other relevant literature).

In this paper, we identify three key issues with the GAModel, and propose adjustments to the algorithms that address these issues.

The first issue is that the genome representation used by GAModel has too many parameters (over 2000 for regular cases). Even though a majority of these parameters do not contribute for the accuracy of the final risk model, the size of the search space implies a slower optimization time. To address this issue, we propose a new genome representation for an earthquake risk model, which we will call "Reduced Representation". In the Reduced Representation, we avoid representing every single location in the area under study, and only those locations with a minimal probability of earthquake are represented as parameters in the evolutionary process. By reducing the search space, this representation is expected to also increase the convergence speed of the evolutionary optimization process.

The second issue is that GAModel does not take into account any sort of domain knowledge, such as the assumption that earthquakes cluster in both time and space. Heuristic search methods such as Genetic Algorithms usually benefit from the introduction of domain knowledge to the search. Therefore, we propose a hybrid version of the GAModel which incorporates seismic models of earthquake decay. This version generates a model with a much smaller number of earthquakes than the regular GAModel. For each earthquake in this model, a sequence of aftershocks is generated using an adaptation of the Epidemic Type Aftershock Sequence model (ETAS). We expect that this hybrid approach will produce more accurate models.

The third issue is the examination of "de-clustering" effects in the historical catalog used for generating the risk model. In seismology, de-clustering refers to the act of identifying earthquakes as either main-shocks or aftershocks, and removing all but the main shocks from the catalog, which is considered the representative earthquake for the group. Accordingly, a de-clustered earthquake catalog is considered to be easier to study, given that the de-clustering process removes redundant information [2]. In this work, we generate the de-clustered catalog by grouping earthquakes in space and time using spectral clustering [3].

These adaptions are described in detail in section 3. We compare the contributions of each adaptation to the generation of models based on the earthquake catalog of the Japanese arquipelago, between 2000 and 2010. The set up of this experiment is described in section 4, and the main results are listed in section 5.

Our results indicate that clustering the earthquake catalog resulted in a significant improvement to the precision of the models generated. On the other hand, the new representation and the hybridization did not seem to improve the results of our models. We discuss the implications of these findings in section 6.

2. Background

An Earthquake Risk model states the probability of earthquake occurrence on a defined area and time period. These models are often based on past ocurrence of earthquakes (historical catalogs). They can also make use of seismic properties of the area under study, such as faults, terrain properties, etc.

The "prediction" of earthquakes is a polemic subject, and no research so far has come close to suggesting that individual large scale earthquakes can be predicted. On the other hand, there is value on the study of earthquake mechanisms and the generation of statistical models of earthquake risk [4].

In our previous work [1], we use a Genetic Algorithm (GA) to optimize an Earthquake Risk Model, which is described in the framework proposed by the Collaboratory for the Study of Earthquake Predictability (CSEP).

In the following subsections we describe both the CSEP framework and the original GAModel. After that, we overview other relevant works combining evolutionary approaches and the study of earthquakes.

2.1. CSEP Forecast Framework

The CSEP framework defines a model in reference to a geographical region and a time period [5]. The geographical region is divided in a grid, where each cell in the grid is called a bin.

For example, in this paper we define the "Kanto" region as as the area covered by latitude N34.8 to N37.05, and longitude E138.8 to E141.05. This area is divided into 2025 bins (a grid of 45x45 squares). Each bin has an area of approximately 25km^2

The model defines a number of expected earthquakes for each bin. This number must be a positive integer. A good model is one where the number of estimated earthquakes in each bin corresponds to the actual number of earthquakes that occurs in that bin during the target time interval.

2.2. The GAModel

Using the CSEP framework described in the previous subsection, we proposed the GAModel [1]. The GAModel uses Genetic Algorithms to generate an earthquake risk model based on earthquake catalog data.

In the GAModel, each individual is a candidade Risk Model, which is equivalent to a prediction $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ in the CSEP framework. The Genetic Algorithm will then select the individuals based on the log-likelihood between the model represented by the individual, and the catalog of earthquake occurrences in the target location and time period.

Genome Representation. In the GAModel, each indian individual is represented as real valued array, where each element in the array is latitude/longitude bin in the target area for the desired model. Each bin is associated to a real value representing the earthquake risk in that location. For the initial population, these values are drawn from an uniform distribution between 0 and 1.

During fitness evaluation, the risk value at each bin is converted to an integer forecast, using a modification of the inverse poisson function depicted in algorith 1. In this algorithm, x is the real value to be converted and μ is the mean of earthquake observations across all bins in the catalog data.

Algorithm 1 Obtain a Poisson deviate from a $[0, 1)$ value

```
$L \leftarrow \exp(-\mu), k \leftarrow 1, prob \leftarrow 1 * x$
while $prob > L$ do
  $k \leftarrow k + 1$
  $prob \leftarrow prob * x$
end while
return $k$
```

Fitness Function. The GAModel uses the log-likelihood value between one individual and the catalog data as its fitness function.

Let an individual $X = \{x_1, x_2, \ldots, x_N\}$, where x_i is the risk value associated with bin i, and N is the total number of bin. From X we obtain the earthquake forecast $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ using algorithm 1. This forecast is also the vector of earthquake quantity expectations.

Let the set of earthquake ocurrences from the catalog be $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$. The calculation of the log-likelihood value for the ω_i observation with a given expectation λ is defined as:

$$L(\omega_i|\lambda_i) = -\lambda_i + \omega_i \log \lambda_i - \log \omega_i! \qquad (1)$$

The joint probability is the product of the likelihood of each bin, so the log likelihood $L(\Omega|\Lambda)$ for the entire vector is the sum of $L(\omega_i|\lambda_i)$ for every bin i:

$$L(\Omega|\Lambda) = \sum_{i=1}^{n} L(\omega_i|\lambda_i) = \sum_{i=1}^{n} -\lambda_i + \omega_i \log \lambda_i - \log \omega_i! \qquad (2)$$

Finally, this value is calculated for each year that composes the training data, using the "time slices" method described in [1].

Evolutionary Operators. GAModel uses One Point Crossover, Polynomial Bounded Mutation, Tournament Selection and Elitism. The chance for using crossover and mutation operators is tested independently for each individual in the new population.

Valuse for the parameters for these operators are listed in Table 1. These values are generally the same as used in [1].

TABLE 1. PARAMETERS USED IN THE GAMODEL

Population Size	500
Generation Number	100
Elite Size	1
Tournament Size	3
Crossover Chance	0.9
Mutation Chance (individual)	0.1
Polynomial Bounded parameters	eta = 1, low = 0, up = 1

2.3. Related Literature

The usage of Evolutionary Computation (EC) in the field of earthquake risk models is somewhat sporadic. Zhang and Wang [6] used Genetic Algorithms to fine tune an Artificial Neural Network (ANN) and used this system to produce a forecast model. Zhou and Zu [7] also proposed a combination of ANN and EC, but their system forecasts only the magnitude parameter of earthquakes. Sadat, in [8], used ANN and GA to predict the magnitude of the earthquakes in North Iran.

There are more works when we discuss EC methods and estimation of parameter values in seismological models. Nicknam et al. [9] simulated some components from a seismogram station and predicted seismograms for other stations. They combined the Empirical Green's Function (EGF) with GA. Kennett and Sambridge [10] used GA and associated teleseisms procedures to determine the Fault Model parameters of an earthquake. By doing so, they demonstrated that non-linear inversion can be achieved for teleseismic problems without any calculation of waves travel times.

Another popular approach is to use EC methots do calculate the Peak Ground Acceleration (PGA) parameter. The works done by Kerh et al. [11], [12] are a combination of ANN and GA to estimate or predict PGA in Taiwan. Their goal was to decide which areas may be considered potentially hazardous areas and they focused on urban areas. They also stated that PGA is inversely proportional to epicentre distance. Cabalar and Cevik [13] work also aimed to predict the PGA, but their work uses genetic programming (GP) and use strong-ground-motion data from Turkey.

Jafarian et al. [14], used GP to develop an empirical predictive equation v_max/a_max ratio of the shallow crustal strong ground motions recorded at free field sites. They found a relation between the v_max/a_max and the earthquake magnitude and the source-to-site distance.

Ramos and Vázques [15] used Genetic Algorithms to decide the location of sensing stations. In this work they achieved, in general, better results with the GA method when compared with the Seismic Alert System (SAS) method and a greedy algorithm method. Saeidian et al. [16] work also based on the same idea of locating sensing stations. They do a comparison in performance between the GA and Bees Algorithm (BA) to decide which of those techniques would perform better when choosing the location of sensing stations. He found out that the GA was faster than the BA.

Huda and Santosa [17] published a paper in which the goal was to find, via GA, the speed of the waves P and S in the mantle and in the earth crust. P waves are indicated as the first fault found in seismological data and S waves are the changes caused in the phase of a P wave [17]. This work aimed to obtain a structure of the Japanese underground.

3. Proposed Changes

In this work, we propose three improvements to the GAModel: A reduced genome representation, Hybridization with the ETAS empirical model, and the clustering of the earthquake catalog. Each of these changes are described below.

3.1. Reduced Genome Representation

In the GAModel, problem is represented as a vector X where each bin corresponds to an element in the vector. As the number of bins in a region numbers into the thousands, this representation leads to a huge search space to be explored.

We have observed that in many cases, the vector of catalog earthquakes is sparse. In other words, most of the elements of X will be zero or close to it. To use this fact to decrease the search space, we propose a “reduced” representation of a risk Model.

A summary of the reduced representation is as follows: First, before initializing the Genetic Algorithm, we estimate the expected total number of earthquakes in the model based on the past data. Then, this value is used as the total number of earthquakes to be added to the model. The reduced representation will be a vector of bin coordinates for each of these earthquakes, representing their position inside the target area. This is much smaller than a representation including each bin as an element.

Implementation. The reduced representation is a vector V of ordered pairs. The first element of this pair is the integer index that identify a bin in the model. The second element of the pair is the number of earthquake occurrences estimated for this bin.

The size of the vector V is calculated as the number of bins in the historical catalog that contain at least one earthquake. For each element in V, the bin index and the estimated number of occurrences are drawn randomly from a uniform distribution.

To generate a model from the reduced representation, we need to go two intermediate steps. The first one is to transform the reduced representation into a regular representation. This is achieved by copying the estimated value of an element to the bin indicated by stored index for that element. Bins that are not indicated by any element in the vector are set to zero estimated earthquakes.

The second step is to apply the inverse Poisson on the estimated values to retrieve the number of earthquakes, as described in algorithm 1.

Operators. The reduced representation can use the same one point crossover as the GAModel, but a different mutation operation is required. The mutation operator works by selecting one element in the vector V, and drawing new values for the index and the estimation parameter from a unform distribution.

3.2. Hybridization with ETAS

The GAModel produces risk models without using any sort of domain knowledge, other than the difference between the individual being evaluated and the earthquake catalog data.

However, one simple observation that could be added to the GAModel is that earthquakes cluster in space and time. Large earthquakes are usually followed by a wave of smaller earthquakes, these pairings being commonly known as *mainshocks* and *aftershocks* [18].

To include this idea into the GAModel, we modify the process which generates a Model from an individual. In this modified process, one individial will only produce mainshocks into the model, afterwards the aftershock are derived from the mainshocks, using empirical seismic laws such as the *modified Omori Law*.

We define this hybridization between empirical seismic laws and the GAModel as the *EMP-GA*. Below, we detail the implementation of both steps.

Implementation. The EMP-GA generates models with mainshocks and aftershocks following a two-step procedure.

In the first step, we use the GAModel to generate a set of mainshock earthquakes, which we will refer to as *synthetic mainshock data*. In the second step, we use seismic empirical equations to obtain the aftershocks from the synthetic mainshock data, and add them to the model.

The process we use to generate aftershocks from the synthetic mainshock data is inspired by the space-time epidemic-type aftershock sequence (ETAS). The total number of earthquakes in a bin is given as

$$\Lambda(t,x,y|\Upsilon_t) = [\mu(x,y)+\sum_{t_i \in t} K(M_i)g(t-t_i)P(x,y)]J(M). \tag{3}$$

In this equation, t is the target time for the model, x, y are latitude/longitude coordinates within the target area, Υ_t is the number of mainshocks derived in the first step, $\mu(x,y)$ is the expected number of earthquakes at the (x,y) bin, t_i is a time interval within t, $K(M_i)$ is the total amount of triggered events, $g(\Delta t)$ is the probability density form of the modified Omori law, $P(x,y)$ is a function that distributes aftershocks in space nearby the mainshock, and $J(M)$ is the ETAS simulated magnitude. Let us explain each of these components below.

Omori’s Law and Triggered Events. The Omori law, which is considered to be an empirical seismic formula which has withstood the test of time [19], [20], is a power law that relates the magnitude of an earthquake with the decay of aftershock activity over time. It can estimate the number of aftershocks based on the mainshocks in the synthetic data generated by one individual in the EMPGA. For this approach, we use the probability density function (PDF) form of the modified Omori law [21], defined as

$$g(\Delta t) = \frac{(p-1)}{c(1+\frac{t}{c})^{-p}}. \tag{4}$$

In this equation, p and c are constants. Utsu [19], summarize the studies of this formula for the Japan case, and describe a range for these variables using the Davidon-Fletcher-Powell optimization procedure. These ranges, used in ETAS, are 0.9 to 1.4 for p, and 0.003 and 0.3 for c.

Also, Δt is the time interval for how long a mainshock may influence or cause an aftershock. According to Yamanaka [22], this value must be choosen carefully, for a value too short will lead to a small number of aftershocks, while a value too long might confound aftershocks and background activity. His work suggests the values $p = 1.3$, $c = 0.003$, and $\Delta t = 30$ days, which we use in this paper.

The total amount of events triggered by a mainshock is represented in equation 3 as $K(M_i)$. To calculate this value, we count the number of aftershocks within a given area A from the mainshock, using the formula

$$K(M_i) = A\ exp([\alpha(M_i - M_c)]). \tag{5}$$

Where $M_c = 3.0$ is the magnitude treshold and $\alpha(M)$ is defined as the inverse of the maginitude, according to Ogata [23]. The area A is obtained using the equation from Yamanaka [22]

$$A = e^{(1.02M-4)}. \tag{6}$$

Using the number of triggered events per magnitude $K(Mi)$, and the Modified Omori PDF $g(t)$, it is possible to calculate the total number of earthquakes generated from a mainshock, by iterating over t_i:

$$\sum_{t_i \in t} K(M_i) g(t - t_i) \tag{7}$$

The resulting aftershocks need to be spread on bins near the mainshock position. The $P(x, y)$ component of equation 3 fills this role. It calculates the position of the aftershocks based on the position of the original mainshock. It simply places each aftershock either north, south, east or west of the mainshock, getting further from the origin after each iteration, until there are no more events to be placed.

Finaly, $J(M)$ is obtained by using the function *etasim*, from the SAPP *R* package [24] that simulates magnitude by Gutenberg-Richter's Law.

The above equations are put together in algorithm 3.2.

Algorithm 2 Aftershock distribution from empirical laws

```
FOR EACH BIN:
if Number of earthquakes in bin > 12 then
  Reduce number of earthquakes in bin to 12
end if
aftershocks = 0
magnitude values for earthquakes in bin = J(M)
for magnitude in magnitudes do
  for t in time do
    aftershocks += g(t)*K(magnitude)
  end for
end for
Use P(x,y) to distribute aftershocks to neighbor bins
```

3.3. Clustering the Catalog Data

The third adaptation to the GAModel that we study in this paper is the clustering of the earthquake catalog data using Spectral Clustering. Unlike the two adaptations described beforehand, this one does not require any change on the algorithm itself, happening instead as a data pre-processing step when building the model. After the catalog data is pre-processed, the Genetic Algorithm is applied normally, using the de-clustered data for fitness evaluation.

This pre-processing step aims to remove redundant information from the catalog, by clustering together earthquakes which are closely related in a mainshock/aftershock relationship [2].

However, because it is difficult to determine exactly when two earthquakes should be clustered together, we choose a non-supervised method, Spectral Clustering, to generate the clusters.

Spectral Clustering involves constructing a similarity matrix of the elements to be clustered, finding the k-Nearest-Neighbours graph (KNN) based on the similarity matrix, calculating the Laplacian matrix of the KNN graph, and performing k-means clustering on the eigenvectors of this matrix.

One of the main characteristics of Spectral Clustering that make it interesting for this problem is that it can be very computationally efficient [25]. This is very important for the clustering of earthquake data, since each data set can contain tens of thousands of earthquakes.

Spectral Clustering Implementation. Let the earthquake catalog data be represented as a vector $X = \{x_1, x_2, \ldots, x_n\ in \Re_d\}$, where n is the number of earthquakes in the catalog, d is the number of attributes that characterize an earthquake in the catalog, and K is the desired number of clusters. The clusters are calculated following algorithm 3.3

Algorithm 3 Spectral Clustering

```
Construct the similarity matrix S
for i in $X$ do
  for j in $X$ do
    if $i$ and $j$ are connected in the KNN graph then
      $s_{i,j} = \exp(-||x_i - x_j||^2/2\sigma^2$
    else
      $s_{i,j} = 0$
    end if
  end for
end for
Matrix $D = nxn$ diagonal matrix where $d_{i,i} = \sum_{j=1}^{n} s_{ij}$
Compute Matrix $L = D - S$
Compute $K$ smallest eigenvectors of $L$
Compute matrix $V = (v_{ij})_{nxK}$, using these eigenvactors as columns.
Compute matrix $U = (u_{ij})_{nxK}$, normalizing the rows of $V$ such as $u_{i,j} = v_{i,j}/\sqrt{\sum_j v_{ij}^2}$
Let each row in $U$ represent a data point, and cluster these points using k-means
for each point $x_i$ in $X$ do
  Assign the cluster of $u_i$ to $x_i$
end for
```

In this algorithm, $||x_i - x_j||$ is the Euclidean distance between data points x_i, x_j. We use the number of nearest neighbors equal to five, and σ, the kernel parameter, equal to 100.

We cluster the earthquakes based on their latitude, longitude, time (in minutes), and depth. By observing the distributions of the eigenvectors, we defined the weight of each dimension in the algorithm:

- latitude and longitude: 150
- time: 7
- Depth: 0.5

4. Experiment Design

4.1. Experiment Design

Claus' Extracts. Statistical analysis: ANOVA on the populational means

Tukey HSD for significant differences on ANOVA

-

Old Text. The first experiment was made to compare the all the models proposed with each other and to discover which method would achieve higher log-likelihood values. We created some scenarios (space/time regions), and we applied the methods for for the regions of Kanto, Kansai, Touhoku and East Japan for a given year (2005-2010) with earthquakes with depth lesser than 100km. We also used 3 kinds of catalogues with the minimum magnitude of 3.0: the JMA and the declustered catalogues form the Window method and the SLC method.

We compared the means of the models log-likelihood values using the ANOVA test. If a group of variables considered for the ANOVA test showed no statistically significant difference, we applied the Paired Student t-test, in the case all groups showed statistically significant difference, the Tukey HSD methodology analysis was used.

The second experiment was made a to compare how the magnitude of the earthquakes influence the models generated. We used the same scenarios from the first experiment. We split the models obtained from these scenarios into slices composed of earthquakes that have magnitude in a given magnitude interval. We calculated the log-likelihood of these slices-models and applied the ANOVA test and the Tukey HSD to compared them.

4.2. Data Sets

The earthquake catalog from the Japanese arquipelago were obtained from the *Japan Meteorological Agency* (JMA) webpage.

The goal of this research is to find existing patterns in the occurrence of earthquakes. For that it is essential to access trustful data and to explore its details. From the *Japan Meteorological Agency* web page we obtained earthquake data about earthquakes in Japan. In this data there are information about earthquakes that happened in or nearby Japan, with the variables: time of the occurrence, magnitude, latitude and longitude and epicentre depth, for the years of 2000 to 2013.

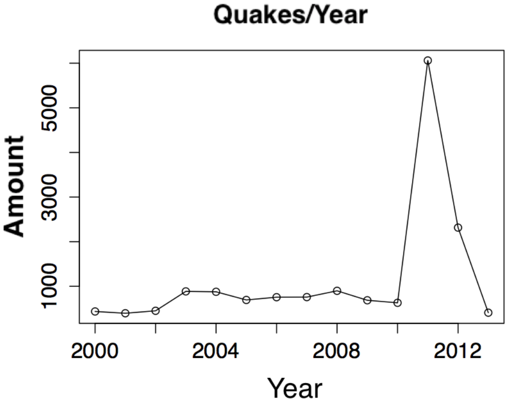

Figure 2. Amount of earthquake by year.

During the preprocessing phase, we discovered a higher number of occurrences of earthquakes during the year of 2011, when a 9.0 M_w earthquake happened, see Section **??**. This earthquake triggered too many after called aftershocks in all Japan. It is considered that big earthquakes may cause others earthquakes [21]. In Figure 2 it is possible to visualise a great number of earthquakes for the year of 2011. Because of this abnormal behaviour and because we decided to focus on more stable occurrences, we limited the training base to earthquakes until 2010.

Based on the statement done before and considering that we want earthquakes that follow more stable patterns, we selected the ones that happened in land areas or very shallow sea areas, with maximum depth of 100km.

For the experiments, the data was changed into slices for every year. Each slice is as follows: if the base contains data about a time interval of 10 years, it will be split in 10 slices.

We also selected some sub-areas in Japan to better extract and understand earthquakes characteristics and patterns. Those areas are Kanto, Kansai, Touhoku and East Japan. The Figure 3 shows how we defined them.

The regions are described as follows:

Kanto Kanto is the region around Tokyo. It is an area with high seismologic activity during the years we studied. Its coordinates are 34.8 North, 138.8 West, with 2025 bins. Each bin covers an area of approximately 25km^2.

Kansai Kansai is the region that includes Kyoto, Osaka and many others historical cities. In this area, rather than Kanto area, there is a small seismic activity. Its coordinates are 34 North, 134.5 West, with 1600 bins. Each bin covers an area of approximately 25km^2.

Touhoku Touhoku is the region in the North of the main Japanese island. It has some clusters of seismic activities during the years we studied. Its coordinates are 37.8 North, 139.8 West, with 800 bins. Each bin covers an area of approximately 100km^2.

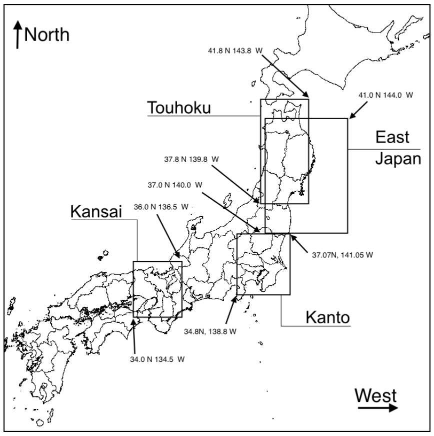

Figure 3. Japan and the areas used in this studied.

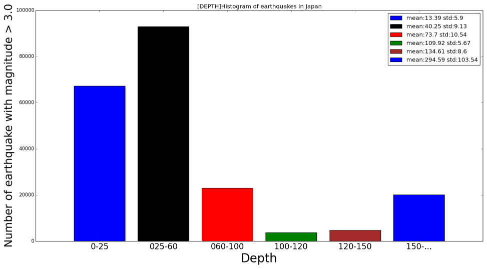

Figure 4. Depth Histogram of earthquakes.

East Japan Is the region that is related with the east coast of Japan. It is the most different area, because it has earthquakes that happened both in land or in the sea. It was in this region that the 9.0 M_w earthquake happened. Its coordinates are 37 North, 140 West, with 1600 bins. Each bin covers an area of approximately 100km^2.

4.2.1. Depth Histogram of Earthquakes. The patterns of earthquakes are dependent of the epicentre. We wanted to explore the relation between the depth of the earthquakes and how would our models behave on those situations.

In Figure 4, it is possible to understand that most of the earthquakes happened with depths smaller or equal to 100 km. The earthquakes deeper than 100 km are fewer and more distant, as it is in the same Figure.

The reason we decided to groups as: earthquakes with depth until 25 km, until 60 km or until 100 km. This is because shallow earthquakes are considered to be more independent earthquakes [22].

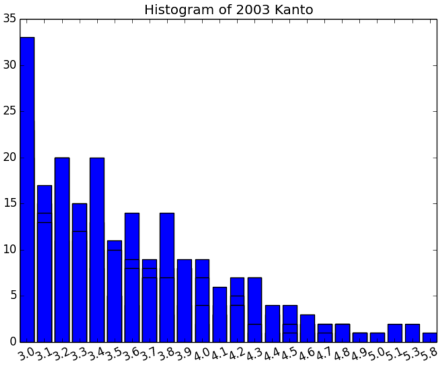

Figure 5. Histogram of earthquakes stronger than 3.0 in Kanto

5. Results

6. Discussion

Acknowledgments

The authors would like to thank the Japan Meteorological Agency for making available the earthquake catalog used in this study.

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*, 2014, pp. 49–58.

[2] T. van Stiphout, J. Zhuang, and D. Marsan, "Seismicity declustering," *Community Online Resource for Statistical Seismicity Analysis*, vol. 10, 2012.

[3] U. V. Luxburg, "A tutorial on spectral clustering," *Statistics Computing*, vol. 17, no. 4, pp. 395–416, 2007.

[4] A. Saegusa, "Japan tries to understand quakes, not predict them," Nature 397, 284, 1999.

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

[6] 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*. IEEE, 2012, pp. 128–131.

[7] F. Zhou and X. Zhu, "Earthquake prediction based on lm-bp neural network," in *Proceedings of the 9th International Symposium on Linear Drives for Industry Applications, Volume 1*, ser. Lecture Notes in Electrical Engineering, X. Liu and Y. Ye, Eds. Springer Berlin Heidelberg, 2014, vol. 270, pp. 13–20. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40618-8_2

[8] 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)*, vol. 109, no. 9, 2015.

[9] 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.*, vol. 119, no. 3, pp. 383–396, June 2010.

[10] B. L. N. Kennet and M. S. Sambridge, “Earthquake location — genetic algorithms for teleseisms,” *Physics of the Earth and Planetary Interiors*, vol. 75, no. 1–3, pp. 103–110, December 1992.

[11] T. Kerh, D. Gunaratnam, and Y. Chan, “Neural computing with genetic algorithm in evaluating potentially hazardous metropolitan areas result from earthquake,” *Neural Comput. Appl.*, vol. 19, no. 4, pp. 521–529, Jun. 2010. [Online]. Available: http://dx.doi.org/10.1007/s00521-009-0301-z

[12] 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*, pp. 1–12, 2015. [Online]. Available: http://dx.doi.org/10.1007/s00521-015-2077-7

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

[14] Y. Jafarian, E. Kermani, and M. H. Baziar, “Empirical predictive model for the v max/a max ratio of strong ground motions using genetic programming,” *Computers & Geosciences*, vol. 36, no. 12, pp. 1523–1531, 2010.

[15] J. I. E. Ramos and R. A. Vázques, “Locating seismic-sense stations through genetic algorithms,” in *Proceedings of the GECCO’11*. Dublin, Ireland: ACM, July 2011, pp. 941–948.

[16] 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.

[17] A. M. Huda and B. Santosa, “Subsurface structure in japan based on p and s waves travel time analysis using genetic algorithm in japan seismological network,” *International Journal of Science and Engineering*, vol. 6, no. 1, 2014. [Online]. Available: http://www.ejournal.undip.ac.id/index.php/ijse/article/view/5762

[18] 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*, vol. 167, no. 8-9, pp. 859–876, 2010.

[19] T. Utsu and Y. Ogata, “The centenary of the omori formula for a decay law of aftershock activity.” *Journal of Physics of the Earth*, vol. 43, no. 1, pp. 1–33, 1995.

[20] F. Omori, “On the after-shocks of earthquakes,” 1895.

[21] J. Zhuang, Y. Ogata, and D. Vere-Jones, “Analyzing earthquake clustering features by using stochastic reconstruction,” *Journal of Geophysical Research: Solid Earth*, vol. 109, no. B5, 2004.

[22] 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*, vol. 38, no. 4, pp. 305–324, 1990.

[23] Y. Ogata and J. Zhuang, “Space–time etas models and an improved extension,” *Tectonophysics*, vol. 413, no. 1, pp. 13–23, 2006.

[24] T. I. of Statistical Mathematics, “Package ‘sapp’,” https://cran.r-project.org/web/packages/SAPP/SAPP.pdf, Jun. 2016, [Online; acessed: 27-07-2016].

[25] X. Ye and T. Sakurai, “Robust similarity measure for spectral clustering based on shared neighbors,” *ETRI Journal*, vol. 38, no. 3, pp. 540–550, 2016.