SEVUCAS: A Novel GIS-Based Machine Learning Software for Seismic Vulnerability Assessment

Saro Lee 1,2, Mahdi Panahi 1, Hamid Reza Pourghasemi 3, Himan Shahabi 4,*, Mohsen Alizadeh 5, Ataollah Shirzadi 6,*, Khabat Khozravi 7, Assefa M. Melesse 8, Mohamad Yekrangnia 9, Fatemeh Rezaie 10, Hamidreza Moeini 11, Binh Thai Pham 12,* and Baharin Bin Ahmad 5

1 Geoscience Platform Research Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124, Gwahak-ro Yuseong-gu, Daejeon 34132, Korea
2 Department of Geophysical Exploration, Korea University of Science and Technology, 217 Gajeong-ro Yuseong-gu, Daejeon 34113, Korea
3 Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz 71441-65186, Iran
4 Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran
5 Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia
6 Department of Rangeland and Watershed Management, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran
7 School of Engineering, University of Guelph, Guelph, Ontario N1G 2W1, Canada
8 Department of Earth and Environment, AHC-5-390, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA
9 Department of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran 16785-136, Iran
10 Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran P.O. Box 19585/466, Iran
11 Organization for Development, Renovation and Equipping Schools of I.R.Iran, Tehran 19338-84883, Iran
12 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
* Correspondence: h.shahabi@uok.ac.ir (H.S.); phamthaibinh2@duytan.edu.vn (B.T.P);
Tel.: +98-91-8665-8739 (H.S.)

Received: 9 July 2019; Accepted: 13 August 2019; Published: 24 August 2019

Abstract: Since it is not possible to determine the exact time of a natural disaster’s occurrence and the amount of physical and financial damage on humans or the environment resulting from their event, decision-makers need to identify areas with potential vulnerability in order to reduce future losses. In this paper, a GIS-based open source software entitled Seismic-Related Vulnerability Calculation Software (SEVUCAS), based on the Step-wise Weight Assessment Ratio Analysis (SWARA) method and geographic information system, has been developed to assess seismic vulnerability by considering four groups of criteria (i.e., geotechnical, structural, socio-economic, and physical distance to needed facilities and away from dangerous facilities). The software was developed in C# language using ArcGIS Engine functions, which provide enhanced visualization as well as user-friendly and automatic software for the seismic vulnerability assessment of buildings. Weighting of the criteria (indicators) and alternatives (sub-indicators) was done using SWARA. Also, two interpolation methods based on a radial basis function (RBF) and teaching–learning-based optimization (TLBO) were used to optimize the weights of the criteria and the classes of each alternative as well. After weighing the criteria and alternatives, the weighted overlay analysis was used to determine the final vulnerability map in the form of contours and statistical data. The difference between this software and similar ones is that people with a low level of knowledge in the area of earthquake crisis management can use it to determine and estimate the seismic vulnerabilities of their houses. This visualized operational
forecasting software provides an applicable tool for both government and people to make quick and correct decisions to determine higher priority structures for seismic retrofitting implementation.

**Keywords:** GIS; SEVUCAS; seismic vulnerability assessment; SW ARA; RBF; TLBO; Tehran; seismic retrofitting

1. Introduction

Earthquakes are considered one of the most destructive natural hazards that cause numerous irreparable physical and financial damage to humans and the environment worldwide [1]. According to the statistics released by the Centre for Research on the Epidemiology of Disasters (CRED, 2015), earthquakes were the third most frequent natural hazard from 1994 to 2013 after flooding and landslide occurrences, resulting in $787 billion of financial damage. Also, during the last two decades, earthquakes and the related disasters, including tsunamis, have been the deadliest natural hazards with a contribution of 55% [2].

Literature reviews [3–5] have shown that the location, magnitude, and exact time of occurrence of earthquakes are still unpredictable. Therefore, disaster management before earthquakes happen plays an important role in reducing earthquake-induced damage to as low as possible. For this purpose, one of the most efficient strategies is to locate vulnerable sites, structures, and infrastructures so their seismic performance can be improved.

In recent years, some researchers have studied the vulnerability of a sample area through investigating the earthquakes that occurred from geotechnical, structural, and socio-economic points of view [6–11]. It is notable that most of these studies have been conducted based on the information from past earthquakes and their relationship with geo-environmental factors. In these studies, multiple-criteria decision-making (MCDM) methods, including analytic hierarchy process (AHP) [9,11–13], fuzzy logic [14–16], fuzzy-AHP [17,18], and technique for order of preference by similarity to ideal solution (TOPSIS) [19–21] in ArcGIS software, have been used. There are multiple studies on the determination of seismic vulnerability that used decision-making methods. However, developing decision-making methods that are capable of easily fulfilling the demands requires expert opinions, and thus, these methods are the main sources of bias and error [22,23]. The possible shortcomings of these decision-making methods can be eliminated using a proper selection of one or a group of evolutionary and intelligent algorithms. Therefore, these sorts of methods can be integrated with various objective algorithms to reduce the error. For this purpose, many researchers have recently tried to apply artificial intelligence algorithms, such as an adaptive neuro fuzzy inference system (ANFIS) [24–26] and evolutionary algorithms [27–30], in geological studies, especially seismic vulnerability assessment, with an emphasis on machine learning algorithms [31].

Although machine learning algorithms are widely used for flood modeling [23,32–40], landslide susceptibility mapping [41–65], land/ground subsidence [51], wildfire [66], sinkhole [67], drought risk [68], gully erosion [69,70], and groundwater vulnerability assessment [22,71–74], they are rarely used in seismic vulnerability assessment. Among the machine learning algorithms, artificial neural network (ANN) is the most widely used method [75], but due to some drawbacks, such as poor prediction power and requiring big data [35], integration with fuzzy logic was implemented, and finally, the ANFIS algorithm was produced. Although the combination of ANN with fuzzy logic and ANFIS is a hybrid method and has a higher prediction power that both ANN and fuzzy logic method, its results and outputs are dependent on the determination of weights in the membership function [76]. Also, literature review has highlighted that ensemble models have a higher performance than individual models, thus hybrid models are strongly recommended.

In the present paper, an integration of three novel and efficient models, involving SWARA, radial basis function (RBF), and teaching–learning-Based optimization (TLBO) methods, was developed
in order to assess seismic vulnerability. According to literature review, these kinds of algorithms are new and do not have the drawbacks of the aforementioned models and it is the first time that an optimization algorithm has been used for seismic vulnerability assessment. Accordingly, new specific GIS-based software entitled Seismic-Related Vulnerability Calculation Software (SEVUCAS) was successfully developed in the C# programming language. The SEVUCAS was tested in our case study in Tehran city in central Iran. The developed software can be used by scientists, analysts, and engineers in different parts of the world to recognize areas vulnerable to earthquakes and provide future hazard mitigation.

2. Description of the Study Area

According to the 2016 census, the population in the megacity of Tehran, the capital of Iran, was more than thirteen million [77] over an area of 730 km$^2$. Tehran is divided into 22 districts, having separate managerial regulations, and is in an area prone to major seismic hazard [78]. This is because they are surrounded by three major and active faults: Mosha, North Tehran, and Rey. Evaluation of the seismic catalogue of Tehran reveals that the majority of historical earthquakes that have affected Tehran have originated from these faults. In this study, the coordinates of the schools in Tehran’s first district is exploited as an explanatory example for SEVUCAS because of the comparatively uniform distribution of population and diversity of structural systems and school buildings, which are used as the disaster management sites after earthquakes. Figure 1 shows the geographical position of the first district of Tehran, the locations of school buildings, and the three aforementioned faults.

![Figure 1. The geographical positions of: (a) District 1 in Tehran and the adjacent faults, and (b) school buildings in District 1.](image)

3. Materials and Methods

After an extensive study and consulting several engineers, seismologists, national managers, and decision-makers, the effective criteria in the four categories of geotechnical, structural, socio-economic, and physical distance were determined.

As stated above, the main purpose of the current study was to develop GIS-based software to provide seismic vulnerability maps for a given area. The proposed algorithm for developing the software requires the following four essential steps that are shown in Figure 2.
3.1. Determining the Main Goal and Criteria

To determine the seismic vulnerability of a considered area, the vulnerability map for each of the four effective criteria—geotechnical, structural, socio-economic, and physical distance—need to be
considered separately. Then, the combination of these vulnerability maps leads to determining the seismic vulnerability of the area. Developing separate vulnerability maps for each of the abovementioned criterion also enabled us to determine the importance of each of them in the seismic vulnerability and thus to facilitate choosing the proper strategy for strengthening and safeguarding the area.

Therefore, the probable damage to any building, i.e., its seismic vulnerability, must be investigated according to its location from different perspectives (geotechnical, structural, socio-economic, and physical distance), in accordance with Panahi et al. [9] and Rezaie et al. [10]. Independent calculation of each vulnerability criterion paves the way for determining its quota in the accumulated vulnerability of a geographical area, and thus, it will be feasible to select a proper risk reduction solution.

In many cases, vulnerability indicators are spatial data; therefore, using the components of ArcGIS Engine can greatly improve the efficiency of the proposed process. In this research, ArcGIS was used as the programming environment. The relevant data for each indicator for all its descriptive features were converted into the ArcGIS software.

3.2. Using the SW ARA Method for Weight Determination

As depicted in Figure 2, SW ARA, which is a new method introduced by Keršuliene [79] for determining the importance and weighting of criteria and alternatives, was used. This method is based on ranking of the factors by the experts [80] (it should be noted that hereafter, “factor” is a general term that refers to criteria as well as alternatives). In order to operationalize this method, the following procedure needed to be taken.

Factors Weighting

For this purpose, the professional viewpoints of experts are used to prioritize factors. Each expert assigns a rank to each factor according to its importance and influence. A weighting is assigned to each of them after different experts have determined their rank. This step is divided into four phases [79]:

- Determining the comparative importance of the average value \(S_j\):
  \[
  S_j = \frac{\sum^n_{i=1} A_i}{n}
  \]
  (1)
  where, \(n\) stands for the number of experts, \(A_i\) shows the offered ranks by the experts for each factor, and \(j\) represents the number of the factor.

- Determining the coefficient \(K_j\) for each factor, as given in Equation (2):
  \[
  K_j = \begin{cases} 
  1 & j = 1 \\
  S_j + 1 & j > 1 
  \end{cases}
  \]
  (2)

- Calculating the recalculated weight \(Q_j\) as follows:
  \[
  Q_j = \frac{K_{j-1}}{K_j}
  \]
  (3)

- Finally, the relative weight of the factor is obtained using the following equation:
  \[
  W_j = \frac{Q_j}{\sum^m_{i=1} Q_i}
  \]
  (4)
  where, \(W_j\) denotes the relative weight of the \(j\)-th factor, and \(m\) represents the total number of factors. A detailed description of SWARA technique can be found in References [81–84].
3.3. Optimization of Each Class’ Weight Values Related to Each Factor

In this study, four main criteria were considered: geotechnical, structural, socio-economic, and physical distance to needed facilities and away from dangerous facilities. Each criterion has some alternatives. For example, in the physical distance to needed facilities criterion: open spaces (OS), road network (RN), police stations (PS), fire stations (FS), etc. are taken into consideration as alternatives in which each alternative has some classes. The effect of the calculated weight using SWARA for each alternative on any of its classes is the same. This equal weight for all values in each class creates two problems: (1) the lower, intermediate, and upper bound of values in the class has the same weight; and (2) weights of two values related to two adjacent classes in the vicinity of their boundaries have considerable differences. To solve these problems, researchers have suggested different solutions [31]. To reduce the effects of abrupt changes of weight values at the boundaries of each class, the proposed method allocates the calculated weight of each class to its intermediate point. In the next phase, by interpolation of the intermediate points and the weight of the classes, the weight of other values for the classes of the factors is determined. Different methods for the interpolation can be used, and among them, intelligent algorithms should be considered. These algorithms, including artificial neural network (ANN), PSO [85], artificial bee colony (ABC) [86], ANFIS [87], and RBF [88], have gained a lot of attraction and are widely used in studies due to their high accuracy compared to other methods. In this study, two intelligent algorithms, namely TLBO and RBF, have been used for interpolating the weights of the considered factors.

3.3.1. Radial Basis Function (RBF) Neural Network

Artificial neural networks are widely used in various fields because of their key feature of training. There are many different networks that are classified based on their structure and training method. One of the most widely used artificial neural networks is RBF, which was first introduced by Broomhead [89]. RBF is a feed-forward artificial neural network that has been widely used in different fields, including geology, mainly due to its high learning speed and having no local minimum problem [90].

Briefly, the RBF structure is made up of three layers; input, non-linear hidden, and linear output layer (Figure 3) [91]. Hidden layers are also made up of a number of nodes, each having a data cluster with a unique influence radius and center. Each node calculates the distance from its own cluster center to the data vector. The distance is calculated using a transformed basis function and is transmitted to outer layer nodes as the hidden layer output [92]. Several functions are used as radial basis functions, which is illustrated in Table 1.

![Figure 3. Structure of an RBF neural network.](image_url)
Table 1. Example of radial basis functions [93].

<table>
<thead>
<tr>
<th>Radial Basis Function</th>
<th>Function Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thin-plate spline (TPS)</td>
<td>$F_{TPS}(x) = x^2 \log(x)$</td>
</tr>
<tr>
<td>Gaussian (G)</td>
<td>$F_G(x) = \exp\left(-\frac{</td>
</tr>
<tr>
<td>Multi-quadrics (MQ)</td>
<td>$F_{MQ}(x) = \sqrt{x^2 + \sigma_i^2}$</td>
</tr>
<tr>
<td>Inverse multi-quadrics (IMQ)</td>
<td>$F_{IMG}(x) = \frac{1}{\sqrt{x^2 + \sigma_i^2}}$</td>
</tr>
</tbody>
</table>

In Table 1, $x = [x_1, x_2, \ldots, x_n]^T$ is the input vector, $F = (F_1(x), F_2(x), \ldots, F_k(x))$ is the $i$-th basis function, $||.||$ is the Euclidian distance, and $v_i$ and $r_i$ are center and width of the Gaussian basic function of the $i$-th node, respectively. Among the basis functions given in Table 1, the Gaussian RBF is the most commonly used basis function [93] because of its appealing properties and flexibility associated with the values of its parameters.

Finally, the output of RBF is calculated using the linear combination set of basis function values, and the weight of each hidden layer node is defined by Equation (5):

$$y_h(x) = \sum_{i=1}^{h} W_{ih}F_i(x)$$

where, $y = [y_1, y_2, \ldots, y_m]^T$ is m-dimensional output vector, $y_h$ is the output of $h$-th node in this vector, $W_{ih}$ is the weight between the $i$-th node of the hidden layer and the $h$-th node of output layer, and $F_i(x)$ is the output of the $i$-th node.

It must be noted that there are different algorithms of the RBF method used by researchers, including the imperialist competitive algorithm (ICA), the teaching–learning-based optimization (TLBO), and the genetic algorithm (GA). However, in all RBF training algorithms, a cost function is used to control the process of training in a way where a low value represents good training and a high value shows poor training. In other words, a high value of the cost function means that the predicted amount by the RBF is far from the target value and a low amount of the cost function shows that the amount predicted by the RBF is close to the target value. The value of the cost function is calculated as follows:

$$E(t) = \frac{1}{2} [y(t) - y_q(t)]^2$$

where $E(t)$ is the cost function at time $t$, and the error $[y(t) - y_q(t)]$ is calculated by subtracting the target value and the value predicted by the RBF at time $t$.

3.3.2. TLBO Algorithm

The other heuristic algorithm used in this study, which has been modelled on the learning process in a classroom, is called the teaching–learning-based optimization (TBLO) algorithm and was introduced by Rao et al. [94]. TBLO, like other evolutionary algorithms, is a population-based method where each part of the population is a learner as well as a solution. This algorithm has become a popular method for the optimization of non-linear issues due to it being a simple and efficient technique [95]. This algorithm consists of two principal phases, teaching and learning, after the creation of the primary population in a phase that is described in the following:
Initialization

In this phase, the initial population \((NP \times D)\) is created randomly, where \(NP\) represents the number of learners and \(D\) indicates the dimension of the instances, i.e., the number of suggested subjects. The initial population is given as the following matrix:

\[
\text{Initial population } x = \begin{bmatrix}
x_{1,1} & \cdots & x_{1,D} \\
\vdots & \ddots & \vdots \\
x_{NP,1} & \cdots & x_{NP,D}
\end{bmatrix}
\]  \hspace{1cm} (7)

Teacher Phase

In this phase, the teacher tries to educate learners and improve the mean value of the class results in two subjects. The best learner, which is the best solution as well, is considered as a teacher.

The mean value of each column represents the mean value of the grades gained by different pupils for each subject is computed using Equation (8) below:

\[
m_d = [m_1, m_2, \ldots, m_D]
\]  \hspace{1cm} (8)

In a specific subject, the result of equivalent teacher and the difference between the mean results are given by Rao et al. [94]:

\[
M_{\text{diff}} = \text{rand}(0,1)[X_{\text{best}} - T_F M_d]
\]  \hspace{1cm} (9)

where \(T_F\) is the teaching factor, \(\text{rand}(0,1)\) is a number between 0 and 1 that is chosen at random. The \(T_F\) value is either 1 or 2 and is decided randomly using Equation (10):

\[
T_F = \text{round}[1 + \text{rand}(0,1)]
\]  \hspace{1cm} (10)

The available population is updated by using Equation (11):

\[
X_{\text{new}} = X + M_{\text{diff}}
\]  \hspace{1cm} (11)

Elements of \(X_{\text{new}}\) are accepted if \((X_{\text{new}}) < f(x)\); else, elements of \(X\) are accepted.

Learner Phase

This is the final phase of the algorithm in TLBO that is used for advancing the local search algorithm capability. In this stage, learners try to use interaction with their classmates in order to upgrade their knowledge. Each learner acquires knowledge from new things that other students choose by chance if the selected students have better knowledge than the learner does. The learner phase is expressed by Equation (12), where two learners \(X_i\) and \(X_j\) \((j \neq i)\) are randomly selected, to give:

\[
X_{\text{new}} = X_i + \text{rand}(0,1)(X_i - X_j), \text{ if } f(X_i) < f(X_j)
\]  \hspace{1cm} (12)

Else:

\[
X_{\text{new}} = X_i + \text{rand}(0,1)(X_j - X_i) \text{ Accept } X_{\text{new}} \text{ if it performs better.}
\]

4. Results and Analysis

After the classification of each factor, the weight of each factor and its class were determined according to the SWARA method. In this method, the upper and lower bounds of each class for factors had a distinct weight; the weight assignment process is improved through the extraction of the weight for all of each factor’s related classes using a combination of TLBO and RBF algorithms, and then the weights were optimized. Finally, components of ArcGIS were employed, making use of the C# programming language to create a spatial environment and the user-friendly software.
4.1. SWARA Weighting

The weighting of the alternatives and criteria are presented in Tables 2 and 3, respectively. Furthermore, the weighting layers are given in Figure 4. According to the SWARA results, in the case of geotechnical, structural, socio-economic, and physical distance to needed facilities and away from dangerous facilities criteria, SWARA values were the highest for peak ground acceleration (0.61), construction materials of structures (0.34), population density (0.28), open spaces (0.25), and gas stations (0.40), respectively. Also, SWARA values were the lowest for slope (0.39), seismic resonance coefficient of structures (0.18), population of employment status density (0.08), disaster management centers (0.08), and high-voltage electrical power transmission lines (0.26) for aforementioned criteria, respectively. On the other hand, from Table 3, it can be seen that geotechnical and physical distance had the maximum and minimum SWARA values, respectively. Additionally, as an example, the weights of each class for the police station alternative obtained by SWARA can be found in Table 4 [10].

Table 2. Calculated weight of alternatives using SWARA.

<table>
<thead>
<tr>
<th>Alternatives/Sub-Indicators</th>
<th>$s_j$</th>
<th>$K_j$</th>
<th>$Q_j$</th>
<th>$W_j$</th>
<th>Description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geotechnical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peak ground acceleration (PGA)</td>
<td>-</td>
<td>1.00</td>
<td>1.00</td>
<td>0.61</td>
<td>Earth movement during earthquakes</td>
<td>[9,96,97]</td>
</tr>
<tr>
<td>Slope</td>
<td>0.55</td>
<td>1.55</td>
<td>0.65</td>
<td>0.39</td>
<td>Surface stability and landslide severity</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4. SEVUCAS input data: (a) peak ground acceleration, (b) slope, (c) construction materials of structures, (d) age of construction, (e) structure quality, (f) seismic resonance coefficient of structures, (g) population density, (h) children population density, (i) elderly population density, (j) female population density, (k) population level of education density, and (l) population employment status density, Distance to: (m) open spaces, (n) road network, (o) police stations, (p) fire stations, (q) hospital, (r) disaster management centers, (s) gas stations, (t) gas pipelines, and (u) high-voltage electrical power transmission lines.
Table 3. Calculated weight of criteria using SWARA.

<table>
<thead>
<tr>
<th>Criteria (Indicators)</th>
<th>$S_j$</th>
<th>$K_j$</th>
<th>$Q_j$</th>
<th>$W_j$</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geotechnical</td>
<td>-</td>
<td>1.00</td>
<td>1.00</td>
<td>0.34</td>
<td>[10]</td>
</tr>
<tr>
<td>Structural</td>
<td>0.20</td>
<td>1.20</td>
<td>0.83</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>Social-economical</td>
<td>0.35</td>
<td>1.35</td>
<td>0.62</td>
<td>0.21</td>
<td>[10]</td>
</tr>
<tr>
<td>Physical distance</td>
<td>0.35</td>
<td>1.35</td>
<td>0.46</td>
<td>0.16</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Each class’ calculated weight of police station as an alternative using SWARA.

<table>
<thead>
<tr>
<th>Class Values</th>
<th>$S_j$</th>
<th>$K_j$</th>
<th>$Q_j$</th>
<th>$W_j$</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1501&lt;</td>
<td>-</td>
<td>1.00</td>
<td>1.00</td>
<td>0.27</td>
<td>[10]</td>
</tr>
<tr>
<td>1201–1500</td>
<td>0.25</td>
<td>1.25</td>
<td>0.80</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>901–1200</td>
<td>0.25</td>
<td>1.25</td>
<td>0.64</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>601–900</td>
<td>0.25</td>
<td>1.25</td>
<td>0.51</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>301–600</td>
<td>0.20</td>
<td>1.20</td>
<td>0.43</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>0–300</td>
<td>0.55</td>
<td>1.55</td>
<td>0.28</td>
<td>0.08</td>
<td>[10]</td>
</tr>
</tbody>
</table>

4.2. RBF and TLBO Algorithms to Optimize Class’ Weight

After weighting each factor using SWARA, these weights were smoothed using a fit-curve method of the medians based on RBF and TLBO (Figure 5). As can be seen, the quality of the weight assignment process was dramatically improved by eliminating the jumps among the classes and in the vicinity of the class’ weights.

Figure 5. Example of the calculated weight using SWARA and its optimization using RBF and TLBO.

4.3. Producing the Final Seismic Vulnerability Map

In this phase, a weighted overlay in the GIS environment was used to overlay the alternatives of each criterion. The final vulnerability map was produced by overlaying four final seismic vulnerability maps derived from the geotechnical, socio-economic, structural, and physical distance criteria for a considered building.

In the SEVUCAS environment, running the model consisted of two separate pages; the input and the output page, which are shown in Figures 6 and 7, respectively. On the input page, all the parameters mentioned in Table 2 were inserted as the input. It was possible to import shape files (*.shp) into SEVUCAS (Figure 6). As represented in Figure 6, after running the software, the final seismic vulnerability map for schools of the Tehran city was obtained. For example, for a selected school
marked in Figure 7, all attributes (PGA = 0.43 gal, slope = 15 degree, etc.) and final vulnerability value (0.174) are visualized. In addition, the vulnerability value of geotechnical, socio-economic, structural, and physical distance criteria for this school were 0.221, 0.164, 0.135, and 0.158, respectively.

![Figure 6. A view of the SEVUCAS input page.](image-url)

![Figure 7. A view of output tab of SEVACUS.](image-url)

In the other words, the geotechnical vulnerability was high for this building due to its vicinity to the North Tehran fault and the 15 degree slope on which the considered building is located. Moreover,
the structural vulnerability was high for this building, given the fact that the building is made of unreinforced masonry and is 44 years of age. This resulted in a high vulnerability in this building class. The socio-economic vulnerability was average for this school since the ratio of women, as well as the ratio of children, to the total population were high. On the other hand, the ratio of elderly people to the population was low and the ratio of educated people to the total population was high; both of which leads to lower vulnerability in this class. Finally, the physical distance vulnerability was low in this case since the considered building was far from the gas pipelines, high-voltage electrical power transmission lines, and the gasoline stations. In general, the resultant vulnerability for this building was high because of high geotechnical and structural vulnerability; both of which possessed a comparatively high weighting factor.

In summary, the results on the output page of SEVUCAS, as given in Figure 7, have four major parts, which are as follows:

(1) The tool could be used to generate geotechnical, structural, socio-economic, and physical distance vulnerability maps in order to separately evaluate the four aforementioned vulnerabilities.

(2) After providing the four vulnerabilities as separate contour maps, the vulnerability of the considered buildings was determined based on their geographical position. Finally, SEVUCAS calculated the overall vulnerability of the considered building by combining all four aforementioned vulnerabilities.

(3) In some cases, a specific building had a great importance such that its seismic vulnerability assessment was included in a case study. The option “Gage” is designed in SEVUCAS to show each of the four abovementioned vulnerability items separately by clicking on that building in the software output window.

(4) In this part, options have been designed to export the numeral vulnerabilities of buildings in the form of Excel worksheets. Moreover, it is possible to sort the considered buildings based on each of the considered vulnerabilities.

5. Discussion

In environmental management, estimating the vulnerability of an area to future earthquakes is highly necessary and significant for decision-making, land management, and the insurance and reinsurance industries. However, unfortunately, there have not been enough attempts to identify vulnerable areas or develop new methods, software, and technology to map them. In this study, we developed a new GIS-based open source software entitled SEVUCAS that is based on the integration of some popular machine learning and expert opinion techniques, involving the SWARA method and geographic information system (GIS), to assess seismic vulnerability. The created software was tested on the group of schools in Tehran, Iran. Since the determination of seismic vulnerability requires exploiting geophysics, civil engineering, urban development, programming, and GIS, this multidisciplinary knowledge required the cooperation of various experts. We developed this new model based on four groups of criteria, namely geotechnical, structural, socio-economic, and physical distance to needed facilities and away from dangerous facilities. Our findings showed that the highest and lowest effective factors in this study area was peak ground acceleration (PGA) (0.61), and population of employment status density and disaster management centers (0.08 for both), respectively, for the modelling of seismic vulnerability (Table 2). The result of our study is consistent with Panahi et al. [9] regarding the seismic vulnerability assessment of school buildings in Tehran city where they indicated that a higher PGA results in a higher probability of liquefaction occurrence is, and hence a higher seismic vulnerability of that area.

A literature review showed that although several methods and techniques have been focused on seismic vulnerability, such as multi-criteria decision-making analysis [18,113] and multivariate statistical analysis [114], few studies have been conducted to develop an optimization algorithm as part of seismic vulnerability GIS-based software. The advantage of this software is the application of the C# programming language using ArcGIS Engine functions to provide enhanced visualization as
well as user-friendly and automatic software for the seismic vulnerability assessment of buildings. Indeed, the RBF and TLBO algorithms enhanced the performance of the SW ARA weights to optimize the weights of the classes of each conditioning factor (Figure 5). In other words, these algorithms increased the performance of SW ARA weights by decreasing the bias and noise between the dataset and the over-fitting problem.

6. Conclusions

In the current research, a hybrid of some novel and efficient models involving SW ARA, radial basis function (RBF), and teaching–learning-based Optimization (TLBO) methods were applied and developed for the assessment of seismic vulnerability and it is for the first time that an optimization algorithm has been used for seismic vulnerability. Finally, a practical GIS-based machine learning software was developed in the C# language programming. In this way, users with a limited level of knowledge in the area of seismic vulnerability can easily determine the vulnerability of their house or office building. The software for computing seismic vulnerability is better at employing powerful spatial- or geographic-related programming tools, e.g., GIS, and used state-of-the-art weighting algorithms, e.g., SW ARA, TLBO, and RBF, as a means to provide the most reliable results. SEVUCAS was developed as a GIS-based open source software with the purpose of making it possible for different users to manually modify the weight factors in order to best suit their case-by-case requirements. Ultimately, the developed software tested using a dataset from Tehran, Iran, and had a reasonable result and showed that the software performs well.

As the PGA and slope stability factors are different from one region to another, the developed GIS-based software needs to be tested and evaluated in other regions to check their applicability and its interpretation. To achieve this goal, this software can be utilized by other researchers around the world to identify earthquake-vulnerable areas, and consequently, mitigate hazards and losses.

Author Contributions: S.L., M.P., H.R.P., H.S., A.S., K.K., A.M.M., M.Y., F.R., H.M., and B.T.P. contributed equally to the work. M.P., H.R.P., H.S., A.S., K.K., A.M.M., M.Y., and H.M wrote the manuscript. S.L., A.M., B.T.P., and B.B.A. provided critical comments in planning this paper and edited the manuscript. All the authors discussed the results and edited the manuscript.

Funding: This research was supported by the Basic Research Project of the Korea Institute of Geoscience, Mineral Resources (KIGAM) funded by the Minister of Science and ICT, and Universiti Teknologi Malaysia (UTM) based on Research University Grant (Q.J130000.2527.17H84).

Conflicts of Interest: The authors declare no conflict of interest.

References

6. Armaș, I. Multi-criteria vulnerability analysis to earthquake hazard of Bucharest, Romania. Nat. Hazards 2012, 63, 1129–1156. [CrossRef]
7. Armas, I.; Gavriș, A. Social vulnerability assessment using spatial multi-criteria analysis (SEVI model) and the Social Vulnerability Index (SoVI model)—A case study for Bucharest, Romania. Nat. Hazards Earth Syst. Sci. 2013, 13, 1481–1499. [CrossRef]


11. Zebardast, E. Constructing a social vulnerability index to earthquake hazards using a hybrid factor analysis and analytic network process (f’anp) model. Nat. Hazards 2013, 65, 1331–1359. [CrossRef]


34. Pham, B.T.; Shirzadi, A.; Bui, D.T.; Prakash, I.; Dholakia, M. A hybrid machine learning ensemble approach based on a Radial Basis Function neural network and Rotation Forest for landslide susceptibility modeling: A case study in the Himalayan area, India. *Int. J. Sediment Res.* **2018**, *33*, 157–170. [CrossRef]


43. Pham, B.T.; Shirzadi, A.; Bui, D.T.; Prakash, I.; Dholakia, M. A hybrid machine learning ensemble approach based on a Radial Basis Function neural network and Rotation Forest for landslide susceptibility modeling: A case study in the Himalayan area, India. *Int. J. Sediment Res.* **2018**, *33*, 157–170. [CrossRef]


64. Shafizadeh-Moghadam, H.; Minaei, M.; Shahabi, H.; Hagenauer, J. Big data in geohazard; pattern mining and large scale analysis of landslides in Iran. Earth Sci. Inform. 2019, 12, 1–17. [CrossRef]


78. Jafarian, M.A. Statistical prediction of the next great earthquake around Tehran, Iran. *J. Geodyn.* 2010, 49, 14–18. [CrossRef]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).