Application of Projection Pursuit Based on Particle Swarm Optimization in Urban Location Vulnerability Assessment

WANG Shao-yu, XIA Quan-wei
School of Management, Harbin Institute of Technology, Harbin 150001, P.R.China

Abstract: Particle Swarm Optimization (PSO) improves Projection Pursuit (PP) in computation amount, and overcomes the disadvantage of many genetic algorithms in operation of "Crossover" and "Mutation". Therefore, this study employs PSO-optimized PP to study the formation of the major location pattern of Harbin, a waterfront city in northern China. It is found that the population living in a specific location pattern is more susceptible to many factors including the terrain and the distance from the water source; the change of urban spatial pattern and the social vulnerability to disasters are in a high coupling relation; and the urban civilians, handicraftsmen and businessmen in Daoli District and Daowai District of low-lying terrain are the populations with the highest social vulnerability, which does not change with time.

Keywords: city, particle swarm optimization, projection pursuit, social vulnerability, spatial pattern

1 Introduction

Since the beginning of the 20th century, natural disasters in China have come to an explosion period in terms of not only type and frequency, but also the loss caused from damage. As disaster causes more and more damages, disaster risk has become one of the greatest obstacles that hinder the sustainable development of human society. With the development of disaster prevention and reduction works to the depth in each countries, both the academic and the industrial fields have turned the attention from the traditional view that took disaster-inducing factors as the core of research to a new one that focuses on the social vulnerability of human society itself. The scholars hold different views on the meaning of social vulnerability: Burton et al. (1978) suggested that it means the susceptibility to the destruction and damage from natural disaster[1]. Timmerman (1981) took it as the degree of the adverse impacts of disaster event on social system[2]. Mitchell (1989) pointed out that social vulnerability is a potential loss[3]. Smith (1992) pointed out that it is the comprehensive metrics of the risk of disaster, and that of the social and economic ability to handle disaster events[4]. Blaikie (1994) believed that it is the ability of individuals or community to predict, handle, withstand and recover from disasters, which involves diverse factors to describe how natural or social disasters threaten our lives[5]. In order to solve the threat of social vulnerability to disaster risk, Blaikie et al. (1994) put forward a disaster pressure and release model[6]. Hewitt et al.(1997) expended the research on social vulnerability to the fields of nature, technology and man-made disasters fields, as well as influencing factors of disaster mitigation[7]. Through the research on social vulnerability, we gradually come to the point that the nature of disaster is the destruction or damage to human beings and human society itself. This potential risk would pose a threat to human society in future to a great extent, that it not only puts great pressure to the current environment of human's survival, but also is likely to cause huge loss to individuals, even the entire society[8]. The research on social vulnerability to natural disasters involves many intricate and complex factors closely linked to human society, such as life safety and social wealth, which are related to the survival of human beings and social development.

2 Research on relation between urban location change and vulnerability and its significance

There is a special formation mechanism of the occurrence and the development of natural disasters. In terms of international cooperation against natural disaster risks, UNESCO clearly mentioned on "Intergovernmental Conference on the Assessment and Mitigation of Earthquake Risk" held by it, that cultural environment, changes in history and economy and social development are important reality reference that influence the higher frequency of natural disasters including earthquake, flood and debris flow in so many developing countries[9]. In 1945, Gilbert F. White pointed out the main principles in understanding the disaster-inducing factors and adjustment to natural extremes; then he combined this concept and the changes of geographical environment for comprehensive consideration, and found that the main cause for the aggravation of flood disaster in this region was the
vulnerability of the human-earth system in this region\textsuperscript{[10]}. Later, Kates et al. mentioned that it was necessary to take into consideration the typical characteristics of social vulnerability for systematic fitting analysis of the types of different natural disasters\textsuperscript{[11]}. In 2000, Downing T. E. put forward that it was necessary to regard social vulnerability as the major research direction of vulnerability science, and concluded that it was necessary to define its fundamental features and research objectives as a research direction\textsuperscript{[12]}. After these researches, the scholars came to a consensus that the traditional research pattern that focused on the inducing factors of natural disasters was broken through; it was required to intensify the consideration of comprehensive indexes, such as the social capability of bearing and adapting to disasters and its adjustment ability; and the research and discussion on social vulnerability under the background of natural disasters should be one of the developing direction of this science.

Disaster science is an empirical research field, in which a classic case is a tracking survey in a small village of Philippine carried out by Allen in 2001. Allen put forward a comprehensive social vulnerability evaluation method taking rural community as basic research unit, which includes leadership style and traditional mores. Another classic case is a research to a U.S. community carried out by Bankoff in 2007, which found that there was a huge difference in values and intervention capacity to natural disasters between different levels (especially social vulnerable levels) due to their difference in financial level\textsuperscript{[13]}. These classic cases tell us from an aspect that comprehensive consideration of the causes for social vulnerability to natural disasters and the social indexes of influencing factors has been included into research scope, and this research will impose important and far-reaching influence on reducing regional disaster risk, increasing anti-disaster ability and recovering ability.

When reviewing the evolution of urban spatial pattern in the West, it is not difficult to find that industrial revolution caused the deep change in the field of socio-economy, brought the West into the process of rapid urbanization, and broke up the traditional urban spatial pattern which had taken family and small industrial park, residential area, business service area and the concept of social aid recipients, namely "slum". The occurrence of this social phenomenon has attracted the attention of some scholars in the field of social improvement, who have discussed this phenomenon, and proposed a research pattern based on the differentiation of social vulnerable groups after the evolution of urban pattern. Moreover, it is worth mentioning that in the middle of 1990's, Cutter concluded the existing researches and put forward the regional disaster model, which introduces the influences of different geographical environments and social backgrounds on natural disaster risk in details. He believed that any change of any factor that is easy to be ignored, such as biocenosis and physical spatial structure, might cause great change of regional vulnerability\textsuperscript{[14]}. In 2004, Dwyer analyzed the disaster-inducing mechanism based on the interaction among the risk of disaster-inducing factor, the extent of exposure to disaster-formative environment and the vulnerability of the disaster-affected region, and put forward a disaster risk evaluation model\textsuperscript{[15]}. In the same year, Smith put forward a disaster risk matrix model based on the contrary relation between physical exposure to natural environment and the vulnerability of human beings\textsuperscript{[16]}. This theory is a response to Castells, who put forward in 1972 the theoretical framework of social system based on spatial structure, political economy and ideology\textsuperscript{[17]}.

3 Research method

3.1 City natural condition

Urban flooding is a chronic disease that recurs almost every year in China, and poses a great threat to the safety of our people's property and lives. This research takes Harbin, an important city in northern China, and Songhua River, a river flooding frequently, as an example for our analysis. Harbin is a frontier city in the north part of Northeast China, and the capital of Heilongjiang Province, with longitude spanning 125°42′–130°10′ E, and latitude 44°04′–46°40′ N. Its administration area consists of 8 districts and 11 counties, with a total land area of 53,000 km\textsuperscript{2}, and a total population of 9,686,100\textsuperscript{[18]}. The area of its city proper is 1,637 km\textsuperscript{2}, in which the urban built-up area is 220 km\textsuperscript{2}. Harbin is the biggest transportation hub in Northeast China, and the biggest inland port in China connecting the Eurasian continent. Geographically, Harbin is mainly in the midstream and downstream of Songhua River, featuring a flat bottom land and marshy environment. For the urban area, the threat of flooding mainly comes from the main stream of Songhua River. The flood before the Harbin section of Songhua River mainly comes from Nen River, Second Songhua River and Lalin River. Thus, Harbin is a typical floodplain city with river hazard. According to history, there were three recorded great floods in the city of Harbin, namely in 1932, 1957 and 1998, with their threats and losses caused shown in Tab.1.

The city proper of Harbin can be divided into three stages of floodplain formed by Songhua River. With an elevation between 118 and 120 m and a flat and low-lying terrain, the first stage includes Daoli District and Daowai District, and is the important business and trade zone in Harbin, as well as the main living area of urban residents. With an elevation between 125 and 155 m and a smooth transition from the first stage but a clear boundary between, the second stage mainly includes
generally with a low level of education), with a total fishers, handicraftsmen, the homeless and the urban poor, proper (where there was the densest population, mainly and Fujiadian (currently Daowai District) as the city Songhua River, taking Butou (currently Daoli District)

spatial pattern can be simply divided into three periods:

The first period is from the end of the 19th century to the time around the construction of the Chinese Eastern Railway. According to historic record, the city of Harbin was established on the low-lying beach of Songhua River, taking Butou (currently Daoli District) and Fujiadian (currently Daowai District) as the city proper (where there was the densest population, mainly fishers, handicraftsmen, the homeless and the urban poor, generally with a low level of education), with a total population of three to five thousand. Since the end of the 19th century, the Imperial Russia sent advance team to the region of Tianjiaishaoguo (currently Xiangfang District) for war and economic factors. After the completion of Eastern Railway, a large amount of white emigres and rich Jewish merchants had moved to Harbin escaping from the punishment of newly-built Soviet Union since the October Revolution. Later they settled down where is Daoli District now, due to the convenient waterway transportation and port opening permission of Qing government, and became the main population of Daoli District. Before Japanese invasion of China, urban spatial pattern of Harbin had come to an early form basically.

The second period was from the establishment of the puppet state of Manchukuo to the early stage of People's Republic of China, when Japan's urban planning was directing the development and evolution of Harbin's urban spatial pattern, which highlighted function partition based on regional structure, planned the residence area, commercial area and major administrative area in ratings strictly, and restricted cross-regional population flow. At the same time, Daoli District on the north side of the railway station extended to the upstream of Songhua River and became its center of economy; the area on the south side of the railway station became the center of politics, where there came a lot of business institutions and administrative organizations, finally forming the new core of the city; and Daoli District extended to the downstream of Songhua River, where the urban civilians mainly lived. This urban pattern was basically kept after new China was founded.

The third period was from the establishment of new China to today, especially along with the reform and opening-up. The change of social environment and the acceleration of urbanization has driven the sharp increase of population density, which has brought new challenges to the development and planning of Harbin. Based on the pattern formed in the second period, this city is speeding up changing the cultural and ecological environment. The economic factors have brought along the formation of a new urban pattern, that more and more landscape view buildings are being built along Songhua River. The establishment of Qunli New District in Daoli District on beach and wetland is a typical example of this period.

According to the history of urban spatial pattern throughout the three periods, and taking into consideration the natural and geographical conditions of Harbin, terrain is a potential factor influencing regional vulnerability.

### 3.2 Formation and evolution of urban spatial pattern

The main areas subject to flooding in Harbin are along the south bank of Songhua River, namely Daoli District and Daowai District which are in lower position, and the population in this area is the main vulnerable group. Historically, the evolution of this area's urban spatial pattern can be simply divided into three periods:

The first period is from the end of the 19th century to the time around the construction of the Chinese Eastern Railway. According to historic record, the city of Harbin was established on the low-lying beach of Songhua River, taking Butou (currently Daoli District) and Fujiadian (currently Daowai District) as the city proper (where there was the densest population, mainly fishers, handicraftsmen, the homeless and the urban poor, generally with a low level of education), with a total population of three to five thousand. Since the end of the 19th century, the Imperial Russia sent advance team to the region of Tianjiaishaoguo (currently Xiangfang District) for war and economic factors. After the completion of Eastern Railway, a large amount of white emigres and rich Jewish merchants had moved to Harbin escaping from the punishment of newly-built Soviet Union since the October Revolution. Later they settled down where is Daoli District now, due to the convenient waterway transportation and port opening permission of Qing government, and became the main population of Daoli District. Before Japanese invasion of China, urban spatial pattern of Harbin had come to an early form basically.

The second period was from the establishment of the puppet state of Manchukuo to the early stage of People's Republic of China, when Japan's urban planning was directing the development and evolution of Harbin's urban spatial pattern, which highlighted function partition based on regional structure, planned the residence area, commercial area and major administrative area in ratings strictly, and restricted cross-regional population flow. At the same time, Daoli District on the north side of the railway station extended to the upstream of Songhua River and became its center of economy; the area on the south side of the railway station became the center of politics, where there came a lot of business institutions and administrative organizations, finally forming the new core of the city; and Daoli District extended to the downstream of Songhua River, where the urban civilians mainly lived. This urban pattern was basically kept after new China was founded.

The third period was from the establishment of new China to today, especially along with the reform and opening-up. The change of social environment and the acceleration of urbanization has driven the sharp increase of population density, which has brought new challenges to the development and planning of Harbin. Based on the pattern formed in the second period, this city is speeding up changing the cultural and ecological environment. The economic factors have brought along the formation of a new urban pattern, that more and more landscape view buildings are being built along Songhua River. The establishment of Qunli New District in Daoli District on beach and wetland is a typical example of this period.

According to the history of urban spatial pattern throughout the three periods, and taking into consideration the natural and geographical conditions of Harbin, terrain is a potential factor influencing regional vulnerability.

### 3.3 Application of evaluation method - PSO-based PP method

#### 3.3.1 Projection pursuit method

For all existing vulnerability evaluation methods, such as gray cluster analysis, fuzzy comprehensive evaluation, matter element analysis and artificial neural network, it is required to determine the grading criteria before evaluation. There are multiple factors influencing the vulnerability to flood disaster, as well as complicated influencing mechanism. Accordingly, there is lack of an approach necessary to determine whether the selected grading criteria is reasonable, and the evaluation results are some discrete vulnerability grades with coarse resolution. Most of the conventional multi-factor comprehensive evaluation methods apply the method of weighted average to combine the influences of these multiple factors as a single index for evaluation, and thus the key becomes to determine the weights. The methods used currently, such as analytic hierarchy process, grey conjunction analysis and Delphi method, tend to be influenced by the subjective factors, and the weights determined by different persons would not be unique. As
a result, it is inevitable that the evaluation results are different. While the exploratory data analysis directly driven the sample data, namely Projection Pursuit (PP) Method, is able to overcome the disadvantages mentioned above.

Accordingly, a new solution has been put forward in international statistic circle in recent years, namely Exploratory Data Analysis (EDA) driven by the data of samples with target evaluation directly originating from samples, in which Projection Pursuit (PP) is one of the typical representatives\textsuperscript{20\cite{21}}. Its main idea is to combine multi-dimensional data at a level, and project the combined configuration to a lower dimensional space for dimension reduction; then the indicator function obtained though projection (or target function) is used to find the optimal solution to the projection indicator function of higher dimensional data based on the structure of indicators exposed during this projection. The steps are shown as below:

Step 1: Nondimensionalizing evaluation indicators. Suppose that the object set is \( \{x^*(i,j) | i = 1\sim n, j = 1\sim p\} \), where, \( x^*(i,j) \) is the value of the \( j \)-th evaluation indicator of the \( i \)-th object; \( n \) and \( p \) are the number of objects and the number of evaluation indicators respectively. There are kinds of evaluation indicators. For the first kind, a greater indicator results to higher vulnerability; and for the other, a smaller indicator results to higher vulnerability. Eq. (1) and Eq. (2) are applicable to these two kinds of evaluation indicators.

\[
x(i,j) = \left\{ \begin{array}{ll}
x^*(i,j) - x_{\text{max}}(j) & \quad \text{for the first kind, a greater indicator results to higher vulnerability} \\
x_{\text{max}}(j) - x^*(i,j) & \quad \text{for the other, a smaller indicator results to higher vulnerability}
\end{array} \right.
\]

Eq. (3)

Step 2: Constructing projection indicator function. PP model is used to combine multi-dimensional data \( \{x(i,j) | j = 1\sim p\} \) to one-dimensional projection \( z(i) \) taking \( a = (a(1), a(2), \ldots, a(p)) \) as projection directions, and classify according to the one-dimensional distribution of \( \{z(i) | i = 1\sim n\} \), where \( a \) is the unit length vector.

\[
z(i) = \sum_{j=1}^{p} a(j)x(i,j)(i = 1\sim n)
\]

Step 3: Optimizing projection indicator function. When the sample set of each indicator is defined, the projection indicator function \( Q(a) \) only changes with the projection direction \( a \). A specific projection direction represents a different data structure characteristic, and the optimal projection direction is the one that is most probable to expose the characteristic structure of high-dimensional data. It is possible to estimate the optimal projection direction by solving the maximization of projection indicator function, and PSO is a common global optimization method, thus it is more effective to use it to solve, which is quite simple and efficient.

\[
\max Q(a) = S_z D_z
\]

\[
s.t. \sum_{j=1}^{p} a^2(j) = 1
\]

Step 4: Clustering. When the optimal projection direction \( a^* \) obtained in the step above is substituted into Eq. (3), the projection \( z^*(i) \) of each sample can be obtained. Apparently, the closer \( z^*(i) \) and \( z^*(j) \) are, the more possible would sample \( i \) and sample \( j \) be classified as the same class. Then \( z^*(i) \) is sorted in descending order, according to which the sample sets can be classified\textsuperscript{22}.

3.3.2 Particle swarm optimization

PP features a huge amount of computation, which limits its further research and the extensive application to a certain extent. Compared to many genetic algorithms which have the disadvantage in the operation of “Crossover” and “Mutation”, Particle Swarm Optimization (PSO) can solve such problems in a simpler, more efficient and quicker manner\textsuperscript{23\cite{24}}. The fundamental of this algorithm originates from the research on the foraging behavior characteristics of bird flock in a region carried out by Dr. Eberhart and Dr. Kennedy in 1995\textsuperscript{25\cite{26}}. It was later widely applied into many research fields since it does not need too many parameters and it is easy to understand and learn, and it has been proven that it is more efficient than genetic algorithms in many cases\textsuperscript{27\cite{28}29}. Its core idea is starting from random solution, and employing principle of moderation to evaluate the quality of such random solution to avoid the limit of function conditions, then applying iteration to find the individual optimal solution (\( p_{\text{best}} \)), so as to find the global optimal solution meeting the preset condition (\( g_{\text{best}} \)). Currently, many scholars have applied PSO algorithm to solve the problem of multi-objective optimization\textsuperscript{30\cite{31}32\cite{33}34}.

\[
v_i = w v_{i-1} + c_1 r_1(p_i - x_i) + c_2 r_2(p_g - x_i)
\]

Where, \( V_i \) is the moving velocity of particle of the current generation; \( V_{i-1} \) the moving velocity of particle of the last generation; \( r_1 \) and \( r_2 \) random numbers,
ranging from 0 to 1; \(c1\) and \(c2\) learning factor, usually \(c1 = c2 = 2\); \(w\) the weight of the influence of the velocity of the last generation on that of the current generation, \(P_i\) is the optimal position of the \(i\)-th particle that has been found so far, and \(P_g\) is the global extremum of the optimal position of entire particle swarm that has been found so far. For the determination of \(w\), \(W_{max}\) and \(W_{min}\) are the weight at the beginning and that at the end respectively; and \(\text{iter}_{max}\) the maximum number of iteration, for which \(\text{iter}\) is the current number of iteration.

\[
W = W_{max} - \frac{W_{max} - W_{min}}{\text{iter}_{max}} \times \text{iter}
\]

During the process of movement, the fitness of each position is to be re-evaluated continually. When the fitness of a \(p_{best}\) is better than \(g_{best}\), its position becomes the optimal particle position \(g_{best}\) in this swarm, so that \(g_{best}\) is refreshed continually until it is the final optimal particle position in this swarm. Moreover, during the movement of the particle swarm, there might be a better fitness of each particle itself, then \(g_{best}\) is refreshed to this position, and the particles are coming close to \(g_{best}\) when moving to \(p_{best}\). Based on the discussion above, this paper combines PP model and PSO (PP-PSO) for the quantitative study of the social vulnerability of each main region in Harbin.

### 4 Data processing and results analysis

Referring to the regional disaster indicators evaluation system put forward by Susan L. Cutter in the middle of 1990's, and taking into consideration the changes of urban locations of Harbin in the recent 100 years, as well as the mechanisms of the formation and evolution of urban floods, this paper selects some representative indicators and regions as listed in the Tab.2.

The PSO-based PP model mentioned above is used to evaluate the vulnerability of the main locations in Harbin. The data of the samples listed in Tab.2 are substituted into Eq. (1) and Eq. (2) respectively to nondimensionalize the indicators, and the results are substituted into Eq. (3), Eq. (5), Eq. (6) and Eq. (4) to get the projection indicator function for evaluation of regional vulnerability. Then PSO is used to optimize Eq. (7) and Eq. (8) to get the values of indicators in projection direction. These values can be regarded as the indicators' contributions to the social vulnerability of the region, which are 0.3865, 0.3541, 0.4837, 0.3125, 0.2439 and 0.5742 respectively. The top two absolute values in the range are the third indicator and the sixth indicator, namely altitude and distance to Songhua River, which have the greatest influence and the second greatest influence respectively. The following two indicators are population density and construction area, and the last two indicators are per capital disposable income and number of schools.

When a value in projection direction is substituted into Eq. (3), the projection result \(z^*(i)\) can be got as shown in Tab.2. Taking the order of projections, the distribution condition and the density of contribution as the criteria, the regions in Harbin can be classified into 3 levels in terms of vulnerability: Level 1 includes high vulnerable regions with projection more than 0.75; Level 2 includes middle vulnerable regions with projection more than 0.45 but less than 0.75; and Level 3 includes low vulnerable regions less than 0.45.

As shown in Tab.2, Daoli District and Daowai District are rated as Level 1, so that they are high vulnerable regions in Harbin, which are also the key regions in terms of flood control. Xiangfang District is rated as Level 2, and Nangang District is rated as Level 3.

### 5 Conclusions

#### 5.1 The change of urban spatial pattern and the social vulnerability to disasters are in a high coupling relationship

Usually, the urban spatial pattern of a city is the result of different factors that are interacting with each

<table>
<thead>
<tr>
<th>S/N</th>
<th>Location Name</th>
<th>Population Density (per km²)</th>
<th>Construction Area (km²)</th>
<th>Distance to Songhua River (km)</th>
<th>Per Capital Disposable Income (RMB)</th>
<th>Number of Schools</th>
<th>Altitude (km)</th>
<th>Value of Projection</th>
<th>Rating of Vulnerability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nangang District</td>
<td>6053.1</td>
<td>60</td>
<td>4.15</td>
<td>25221</td>
<td>70</td>
<td>140</td>
<td>0.1738(4)</td>
<td>III</td>
</tr>
<tr>
<td>2</td>
<td>Daoli District</td>
<td>1482.2</td>
<td>22.6</td>
<td>0.12</td>
<td>22430</td>
<td>42</td>
<td>119</td>
<td>1.1459(2)</td>
<td>I</td>
</tr>
<tr>
<td>3</td>
<td>Daowai District</td>
<td>2770.5</td>
<td>45.73</td>
<td>0.11</td>
<td>13067</td>
<td>39</td>
<td>118</td>
<td>1.5642(1)</td>
<td>I</td>
</tr>
<tr>
<td>4</td>
<td>Xiangfang District</td>
<td>2219.4</td>
<td>29.5</td>
<td>6.17</td>
<td>18012</td>
<td>58</td>
<td>160</td>
<td>0.4824(3)</td>
<td>II</td>
</tr>
</tbody>
</table>

other. Our analysis on the spatial evolution of each main region of Harbin, especially Daoli District and Daowai District, shows that the special history, the social culture and the economic interest during urbanization are the main causes for the current pattern of the region, as well as its evolution. The first stage, which mainly consists of Daoli District and Daowai District, is the high-exposure area in flooding, and has a higher social vulnerability. Therefore, the evolution of spatial pattern of Daoli District and Daowai District and the social vulnerability of these areas are a high coupling relationship.

5.2 The social vulnerable groups do not change significantly with time.

According to the history of Harbin in about a hundred years, and the formation of the urban spatial pattern of this city, the closer to Songhua River the area is, the higher risk of flooding it faces, especially Daoli District and Daowai District. For Daowai District, the main populations under the threat of flooding are the urban poor, handicraftsmen, freelancers and the homeless in so many years, who are in lower class with less income and fewer educational resources, and have limited social recovery and self-building capability. For Daoli District, which is the main business zone of Harbin, the main populations under the threat of flooding are merchants, who have adequate recovery capability, however, their social vulnerability is still great as they are on the stage of terrain which is in the special terrain exposure zone. The data indicate that the damaged groups in high vulnerable regions in these three floods have not changed with time, and these groups are still in high risk.

References


