

ORIGINAL RESEARCH

Modeling the warning system of tour sustainable development with gravitational search algorithm support vector regression

Jing Leng*

Institute of Electrical Engineering, Yanshan University, Qinhuangdao, China

Received: April 16, 2014
DOI: 10.5430/air.v3n2p34

Accepted: May 28, 2014
URL: <http://dx.doi.org/10.5430/air.v3n2p34>

Online Published: June 24, 2014

Abstract

In this paper, a warning system is constructed using Gravitational Search Algorithm Support Vector Regression (GSA-SVR). The gravitational search algorithm (GSA) is used to optimize the regularization parameter of Support Vector Regression (SVR) and is compared to particle swarm optimization. First, the history data of each index are normalized to (0,1). Then, the weights of each index are determined by using grey relationship theory, and meanwhile, the degrees of sustainable development of each year are calculated. The sustainable development of each year is used as sample to train the SVR. The trained SVR is utilized as the warning model to predict the degree of sustainable development in future years. The proposed method is applied to the early warning of tour sustainable development of Qinhuangdao. The simulation results show the effectiveness of the proposed method.

Key Words: Tour sustainable development, Support vector machines, Gravitational search algorithm

1 Introduction

Qinhuangdao is a famous seaside resort which is located in the Bohai Sea. With a pleasant climate and beautiful scenery, it has the tremendous potential for development to become a fascinating tourist city. There are numerous tourist resources in Qinhuangdao—the territory of the Great Wall in Qinhuangdao, the Beidaihe summer resort, the historic city of Shanhaiguan, the Nandaihe beach as well as the Gold Coast in Changli.^[1] What is more, Beidaihe is one of the four major bird-watching resorts in the world.

Tourism is an important industry of the national economy in Qinhuangdao City. Health and sustainable development of tourism plays a significant role in promoting the economic development in Qinhuangdao. In recent years, tourism of Qinhuangdao has a rapid development in order to meet the increasing demand from people out of the city. With the growing number of visitors, environment of scenic surrounding has been damaged in varying degrees. And ex-

ceeding tourists exert great pressure on social infrastructure (such as transportation, accommodation, etc.) and social resources, some areas are even over-exploitation of tourism resources. All of these effects constrain the health and sustainable development of tourism in Qinhuangdao.

Therefore, it is essential to build an early warning system for prediction for the future sustainable development of tourism in Qinhuangdao city. And that is very meaningful to predict and forecast based on predictions of spatial and temporal scales, so that we are capable to come up with some effective solutions.^[2] Generally speaking, these kinds of warnings are whole processes which include discovery and analysis the warning signs, searching for sources, judging the degree and taking right approaches to exclude such dangers. And the early warning system in tourism contains these processes.

Sustainable tourism development of early warning system is based on tourism at present to predict the future status. Sub-

*Correspondence: Jing Leng; Email: liudanyanda@gmail.com; Address: Institute of Electrical Engineering, Yanshan University, Qinhuangdao, China

sequently, it takes corresponding measures according to that status. At present, sustainable development of early warning has attracted much attention among many scholars and they have proposed corresponding methods like BP neural network, composite index, fuzzy comprehensive evaluation method, matter-element analysis, etc. For Example, reference 3 used decision tree to build an early warning model of Huposhuihua in Beijing. In ref. 4, the neural network method was employed to predict the degree of sustainable development in a certain area in Ning Bo; Xiu. P. Y used fuzzy BP neural network to construct tourism sustainable system in Shi Jia Zhuang.

Neural network can approximate any consecutive non-linear function defined on support, and can be used for time series prediction. However, the traditional BP neural network contains multi-layer, so the structure is rather complicate. Besides, BP is essentially a gradient algorithm, so it is easy to fall into local optimum. Support vector machines have been introduced as a powerful robust approach to classification and regression problems. Support vector regression (SVR) is based on Vapnik's ϵ -insensitive loss function and structural risk minimization. It aims to bound the mean approximation error.^[4] In 1995, Vapnik and later Smola explored and developed the SVM approach for regression problems. In support vector regression, a nonlinear model is represented by an expansion in terms of nonlinear mappings of the model input. The nonlinear mappings define a feature space, which may have infinite dimensions.^[5] The solution, however, can be expressed in terms of an associated kernel which has the same dimension as the data set, and it is therefore not necessary to compute the nonlinear mappings explicitly. This type of machine is typically formulated using quadratic optimization under the umbrella of convex optimization. Support vector regression has several appealing properties for black-box identification.^[6]

Sustainable development of tourism is a very complex process, which involves many areas. By extensively searching information and asking for experts, we make a reasonable classification and conclusion about early warning indicators for sustainable development of tourism, and establish a clear system. Take as a reference and in accordance with the characteristics of Qinhuangdao, this paper chooses four types of representative indicators, including indicators of economic development tourism, eco-environmental support index, source index of market development capabilities. Each class also includes a number of the number of secondary indicators.

1.1 Data normalization

There is a great number of early warning indicators of sustainable development, and they have different properties, dimension and order of magnitude. Therefore, in order to avoid to be affected by these factors and ensure objectivity, the original data of indicators are normalized into the

range from 0 to 1. There are lots of methods of normalization, such as: range transformation,^[7] the relative number method.^[2] In this paper, we use the range transformation to normalize the data of early-warning indicators. The idea of range transformation is to make data range standardization transformation, and use the value of each record value minus the minimum value, and then divided by the value of the range record of formal. The formula is as follows:

$$x'_{ij} = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \quad (1)$$

Where x_{ij} denotes the i -th index in the j -th year data values, $\min x_{ij}$ is the i -th index statistics over the years the minimum, $\max x_{ij}$ is the i -th index statistics calendar year maximum.

1.2 Determine the weight of index based on grey relational theory

Determine the weight of each indicator is an important step of calculating the degree of sustainable development. Generally, there are two types of methods to determine the weight. The first category is the subjective weighting method, such as Delphi method,^[8] analytic hierarchy process (AHP)^[9] and so on. The second category is the objective weighting method, such as entropy weighting method,^[10] the principal component method,^[11] and cluster analysis.^[12] Sustainable tourism development of early warning system is a rather complex open system which influenced by many factors interacted with each other. Since rely on subjective opinions of experts is often difficult given the scientific importance of each indicator, in this paper, the use of gray correlation method to determine the weight of each indicator. Grey relational analysis (GRA) is a multi-factor statistical analysis method, which is based on sample data of each factor according to gray correlation factor to describe the relationship between the strength, size and sequence.^[13-15] So that we can determine which are the main factors-large impact on system development that should be given to greater weight; and which is a secondary factor-less impact on system development that should be given to smaller weight.

GRA method assumes that there is a gray system including a set of reference sequence $x_0(k)$, $k = 0, 1, \dots, N$ and a set of referred sequences are $x_i(k)$, $k = 0, 1, \dots, N$. The reference sequence and referred sequences are specific correlation calculated as follows:

- 1) Non-dimensional data processing described in this step as 2.2;
- 2) Find absolute difference between referred sequence and reference sequence.

$$\Delta x_i(k) = |x_0(k) - x_i(k)|, i = 1, \dots, M; k = 1, \dots, N \quad (2)$$

3) Find the maximum and the minimum in sequence.

$$P = \max \max \Delta x_i(k), Q = \min \min \Delta x_i(k) \quad (3)$$

4) Calculating coefficient correlation:

$$\xi_i(k) = \frac{Q + \rho P}{\Delta x_i(k) + r h o P} \quad (4)$$

5) Find the correlation:

$$r_i = \sum_{k=1}^N \xi_i(k) \quad (5)$$

For each indicator, its weight is:

$$w_i = \frac{r_i}{\sum_{i=1}^L r_i} \quad (6)$$

After determining the indicators and the corresponding weight, the degree of sustainable development in a particular year is calculated by following the formula

$$S = \sum_{i=1}^L w_i x_i \quad (7)$$

$$\max L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(X_i, X_j), s.t. \sum_{i=1}^l \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, i = 1, 2, \dots, l \quad (9)$$

The decision function can be shown as:

$$f(X) = \text{sign}[\sum_{i=1}^l y_i \alpha_i K(X_i, X) + b] \quad (10)$$

The most common kernel functions used in SVM are shown as follows:

Linear kernel

$$K(x, x_i) = \langle x \bullet x_i \rangle$$

Polynomial kernel

$$K(x, x_i) = (\langle x \bullet x_i \rangle + C)^d$$

RBF kernel

$$K(x, x_i) = \exp(-\|x \bullet x_i\|^2 / 2\sigma^2)$$

In this study we focus on the radial basis function kernel for its good performance and universal application.

2.2 Train SVR using particle swarm optimization and gravitational search algorithm

2.2.1 Particle swarm optimization

PSO is an evolutionary optimization technique developed by Kennedy and Eberhart in 1995 inspired by the behavior of

According to 2005-2009 data from Bureau of Qinhuangdao City, table 1 lists the various indicators of sustainable tourism development and their standardized values, where their weights are calculated by Grey relational grade.

2 Warning system based on SVR

2.1 SVR

In this section, SVR is briefly reviewed. For more details about SVR, one can refer to.^[4-6] The basic idea of SVM is mapping the training samples from the input space into a higher dimensional feature space via a mapping function ϕ . Given a training set $X = (x_i, y_i), i = 1, \dots, l$, where x_i are the input vectors and y_i the labels of the x_i , the target quadratic programming problem is

$$\min \Phi(w) = \frac{1}{2} \langle w \bullet w \rangle + C \sum_{i=1}^l \xi_i \quad (8)$$

where C is a penalty parameter, ξ_k are non-negative slack variables.

So the problem of constructing the optimal hyperplane is transformed into the following quadratic programming problem:

bird flocking and fish schooling.^[16] The basic PSO model consists of a swarm of particles moving in a d-dimensional search space where a certain quality measure, the fitness, can be calculated. Each particle has a position represented by a vector $X_i = (x_i^1, x_i^2, \dots, x_i^d)$, and a velocity represented by vector $V_i = (v_i^1, v_i^2, \dots, v_i^d)$, where i is the index of particle. While searching optimal solution in search space, each particle remembers two variables: one is the best position found by its own so far, denoted by p_i and another is the best position among all particles in the swarm, denoted by p_g . As time passes, each particle updates its position and velocity to a new value according to formula (11) and (12)

$$v_i(t+1) = w * v_i(t) + c_1 * r_1 * (p_i - x_i(t)) + c_2 * r_2 * (p_g - x_i(t)) \quad (11)$$

$$x_i(t+1) = x_i(t) + v_i(t+1), i = 1, 2, \dots, n \quad (12)$$

Where w is the inertia weight, c_1 and c_2 are the learning factors, generally, $c_1 = c_2 = 2$, r_1 , and r_2 are two random

number within the interval of (0,1), t is the number of iteration. Generally, the value of each components in v_i can be clamped to the range $[-v_{max}, v_{max}]$ to control excessive roaming of particles outside the search space. The particle flies toward to a new position according to (4). This process is repeated until a user-defined stopping criterion is reached.

2.2.2 Gravitational search algorithm

The Gravitational Search Algorithm (GSA) was first introduced by Rashedi as a new stochastic population-based heuristic optimization tool. It is one of the newest heuristic algorithms that has been inspired by the physical laws is GSA (Gravitational Search Algorithm). This heuristic algorithm has been inspired by the Newtonian laws of gravity and motion.^[17]

To describe the GSA consider a system with s masses in which position of the i th mass is defined as follows:

$$X_i = (x_i^1, x_i^2, \dots, x_i^n), i = 1, 2, \dots, N \quad (13)$$

Where x_i^d is the position of the i th mass in the d th dimension and n is dimension of the search space.

Mass of each agent is calculated after computing current

population's fitness as follows:

$$q_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \quad (14)$$

$$M_i(t) = \frac{q_i(t)}{\sum_{j=1}^s q_j(t)} \quad (15)$$

where $M_i(t)$ and $fit_i(t)$ represent the mass and the fitness value of the agent i at t , respectively. For a minimization problem, $word(t)$ and $best(t)$ are defined as follows:

$$best(t) = \min_{j \in (1, \dots, s)} fit_j(t) \quad (16)$$

$$worst(t) = \min_{j \in (1, \dots, s)} fit_j(t) \quad (17)$$

To compute acceleration of an agent, total forces from a set of heavier masses that apply on an agent should be considered based on the law of gravity (Eq. (16)), which is followed by calculation of agent acceleration using the law of motion (Eq. (17)). Afterward, velocity and then position of an agent are updated according to Eqs. (18) and (19):

$$F_i^d(t) = \sum_{j \in k, bestj \neq i} rand_j G(t) \frac{M_j(t)M_i(t)}{R_{ij}(t) + \varepsilon} (x_j^d(t) - x_i^d(t)) \quad (18)$$

$$a_i^d(t) = \sum_{j \in k, bestj \neq i} rand_j G(t) \frac{M_j(t)}{R_{ij}(t) + \varepsilon} (x_j^d(t) - x_i^d(t)) \quad (19)$$

$$v_i^d(t+1) = rand_i \times v_i^d(t) + a_i^d(t) \quad (20)$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \quad (21)$$

where $rand_i$ and $rand_j$ are two uniformly distributed random numbers in the interval $[0,1]$, ε is a small value, $r_{ij}(t)$ is the Euclidean distance between two agents i and j , defined as $\|x_i(t), x_j(t)\|_2$. $kbest$ is the set of first K agents with the best fitness value and biggest mass, which is a function of time, initialized to $K0$ at the beginning and decreasing with time. Here $K0$ is set to s (total number of agents) and is decreased linearly to 1.

The gravitational constant, G , is a decreasing function of time where it is set to $G0$ at the beginning and is decreased exponentially towards zero at the last iteration.

$$G = G_0 \exp\left(-\frac{\alpha \times t}{T}\right) \quad (22)$$

where T is the total number of iterations. The pseudo-code of the RGSA is given by Fig. 1. The GAS algorithm is composed of the following steps:

- (a) Searching space identification.
- (b) Randomized initialization.
- (c) Fitness evaluation of agents.
- (d) Updating $G(t)$, $best(t)$, $worst(t)$ and $M_i(t)$ for $i = \frac{1}{4} 1, 2, \dots, N$.
- (e) Calculation of the total force in different directions.
- (f) Calculation of acceleration and velocity.
- (g) Updating agents' position.
- (h) Repeat steps (c) to (g) until the stop criteria is reached.
- (i) End.

2.2.3 The training procedure of SVR using GSA

The SVM parameters (γ , C), and feature mask are chosen as the position of super masses of the GSA. The training procedure of GSA-SVM is given in Fig. 1.

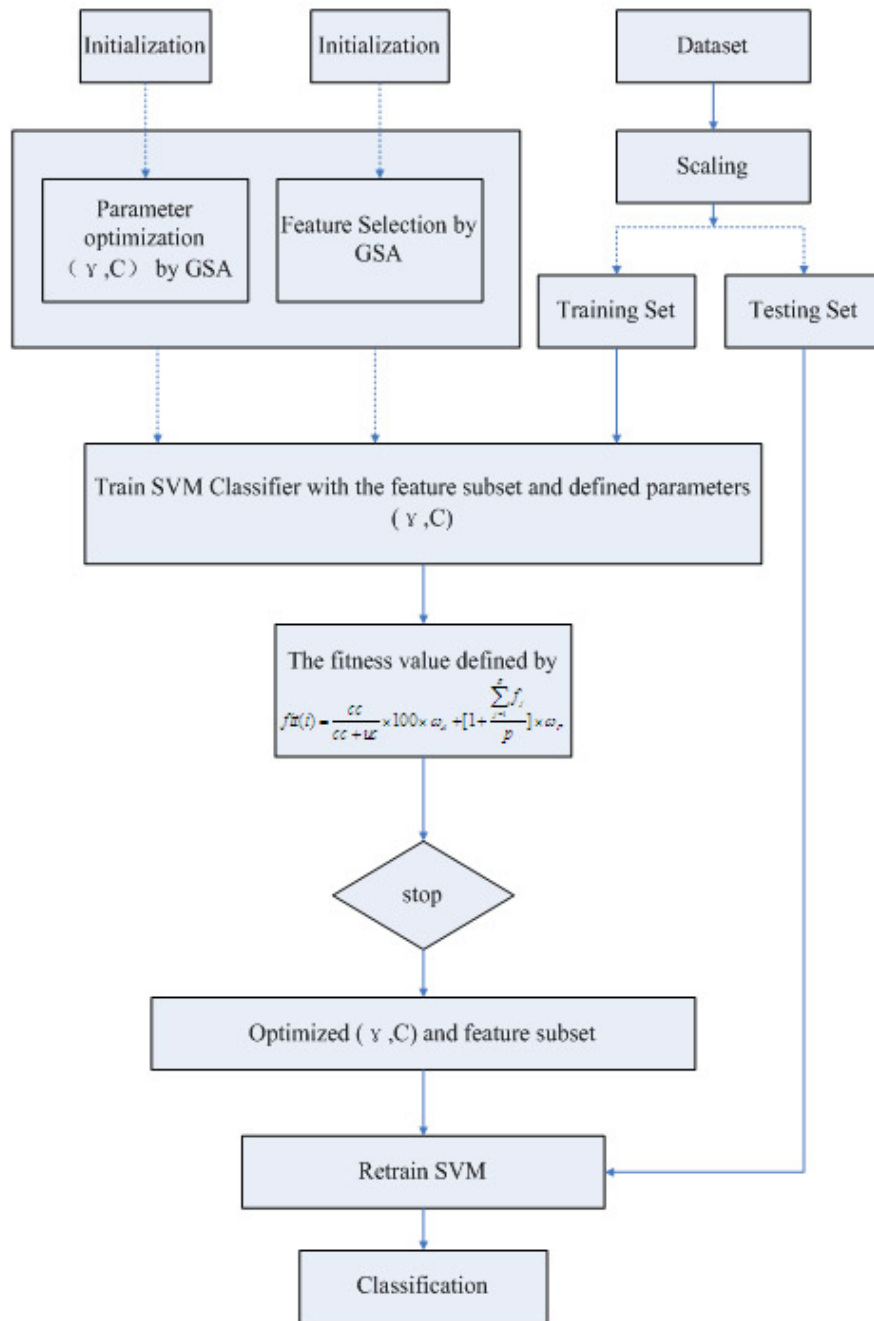


Figure 1: The training procedure of GSA-SVM

2.3 The establishment of early warning models

Take the statistical data of individual indicators in table 1 as the training sample of SVR. Finally, the trained SVR is obtained as an early warning model.

2.4 Classification of early warning

In order to determine the extent of the crisis of warning, the warnings are classified into three to five additional warning limits, with different colors of lights on behalf of crisis lev-

els of warning. The methods of determining class and degree of warning include systematic method of classification, control chart method, catastrophe theory method, contrast judgments, expert decision and comprehensive evaluation to determine the value. This article evaluates the value of tourism in sustainable development, which divided into five intervals between 0 and 1, corresponding to the outputs of red, orange, yellow, green and blue which represent weak sustainability, basic sustainability, normal sustainable, fair sustainable and strong sustainable, respectively. As shown in Table 2.

Table 1: Early warning indicators of sustainable tourism development and standardization of the value table of qinhuangdao city

Top indicator	Secondary indicator	Weigh	2005	2006	2007	2008	2009	2010	2011	2012
Tourism Economic Development Indicator	GDP	0.0535	0.6004	0.6604	0.6507	0.672	0.6731	0.6725	0.6689	0.6705
	Total tourism income	0.0608	0.8654	0.8524	0.8647	0.8659	0.8624	0.8655	0.8621	0.8654
	Total tourism revenue growth	0.0635	0.7325	0.8452	0.7326	0.7341	0.7314	0.7337	0.7316	0.7401
	Foreign Exchange Earnings	0.0608	0.6987	0.8465	0.6979	0.6985	0.7001	0.7004	0.7011	0.7106
	Growth of tourism foreign exchange earnings	0.0383	0.8587	0.8536	0.8591	0.8584	0.8594	0.9000	0.8805	0.7562
Ecology Environment Support Indicator	Tourist area / total area of destination	0.0434	0.6601	0.6604	0.6607	0.6611	0.6617	0.6623	0.6631	0.6658
	Environmental damage events	0.0397	0.8632	0.8524	0.8969	0.9011	0.905	0.9532	0.9601	0.9623
	Waste volume / number of tourists	0.0399	0.8324	0.8452	0.8346	0.8325	0.8583	0.8259	0.8419	0.8561
	Whether the environmental protection measures	0.0404	0.8252	0.8465	0.8586	0.8589	0.8624	0.7685	0.8291	0.8419
	Forest cover	0.0446	0.8325	0.8536	0.8598	0.9621	0.8765	0.8974	0.8755	0.8851
	Sewage treatment rate	0.0446	0.8241	0.8256	0.8321	0.8423	0.8513	0.8601	0.8622	0.8654
	Environmental investment / expenditure	0.0526	0.6154	0.6225	0.6347	0.642	0.6724	0.6658	0.6745	0.6782
Source Market Open up Ability Indicator	Tourists / tourist area	0.0498	0.7456	0.7571	0.7579	0.7456	0.7589	0.7591	0.7611	0.7642
	Well-known outside the region	0.0455	0.8463	0.8468	0.8467	0.8453	0.8504	0.8509	0.8543	0.8555
	Significant degree of overall image	0.0464	0.778	0.8024	0.8038	0.765	0.8054	0.8105	0.8137	0.8162
	Visitor satisfaction	0.0455	0.7023	0.7102	0.7203	0.7127	0.7201	0.7301	0.7364	0.7371
Community Culture Progress Indicator	The number of accidents	0.0462	0.7652	0.7358	0.7435	0.7248	0.7456	0.7698	0.7733	0.7801
	Development and protection of tourism resources	0.0464	0.9054	0.9024	0.9109	0.9004	0.9211	0.9119	0.9106	0.9099
	Conflict with the residents	0.0464	0.897	0.8765	0.875	0.81	0.8811	0.8823	0.8839	0.8853
	Characteristics of tourism products	0.0455	0.5684	0.5871	0.5641	0.589	0.5901	0.5921	0.5934	0.5942
	Culture construction	0.0455	0.6325	0.6329	0.6337	0.6401	0.6405	0.6413	0.6435	0.6464
Sustainable Development Degree -S			0.7596	0.7651	0.7697	0.7694	0.7772	0.7785	0.7813	0.7816

Table 2: Valuation form sustainable development of tourism

Evaluation value	[0,0.2)	[0.2,0.4)	[0.4,0.6)	[0.6,0.8)	[0.8,1]
Sustainable development	weak sustainability	basic sustainability	normal sustainable	fair sustainable	strong sustainable
signal	red	orange	yellow	green	blue

3 Prediction of sustainable development in QINHUANGDAO City, 2011

An early warning system for the sustainable development of Qinhuangdao City is developed based on the SVR. Applica-

tion of SVR includes three parts, i.e., the training process, the testing process and the forecasting process, which are stated as follows:

1) The training process: the data from 2005 to 2010 is used

as samples to train SVR, the data of the 21 indicators are used as the input of SVR, the 21 indicators in 2011 as the SVR's output data to form the training set. GSA is used to train SVR.

2) The test process: According to the trained SVR, the data from 2006 to 2011 are used as samples of SVR, we can obtain the data of 21 indicators in 2012. Thus we can calculate the degree of sustainable development in 2012 is 0.7814 when using gravitational search algorithm while 0.7828 using particle swarm optimization. It is shown that gravitational search algorithm has high accuracy. So we use it to predict in 2013.

3) The prediction process: the data from 2007 to 2012 are used as the input to predict the output of each indicator in 2013. We get the predicted value in 2013, resulted in sustainable development degree of 0.7808.

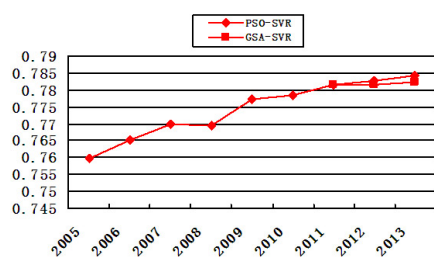


Figure 2: Warning sustainable development of tourism trend comparison chart

Figure 2 shows the Qinhuangdao City from 2005 to 2011 tourism trends for sustainable development of early warning. We can see that the overall upward trend in from 2005 to 2011 degrees which are in the green zone as a strong sustainable development. While in 2012-2013, the growth trend is slow down.

4 Conclusion

A warning system is constructed by using GSA-SVR. The proposed method is applied to building the early warning model of tour sustainable development of Qinhuangdao. Comparison of the PSO-SVR method and GSA-SVR method shows the latter performs better. Based on the statistics in recent years, studies have shown that the sustainable development of the tourist city of Qinhuangdao refers to a strong degree, but early warning analysis is largely empirical analysis and a number of practical factors are not taken into account. Therefore, we can no longer give considerable attention to sustainable development just based on the sustainable development degree. Additionally, to realize a wholesome sustainable development, each of us needs to firmly establish the awareness of risk and awareness of early warning, and protect our environment. Meanwhile the government should strengthen macro-control.

References

- [1] Statistical Yearbook of Qinhuangdao (in Chinese). 2010, pp. 423-424
- [2] X. Yang, G. Weng and Q. Zhang, Research of early-warning system about tourism sustainable development based on chaotic BP neural network, *Future and Development*. 2011; 4: 25-28.
- [3] Y. Zeng, Z. Yang and J. Liu. Algalbloom prediction models for Liuhai-lake in Beijing city, *Advances in Water Science*. 2007; 18: 79-85.
- [4] Zhang X L, Chen X F, He Z J. An ACO-based algorithm for parameter optimization of support vector machines. *Expert Systems with Applications*. 2010; 37(9): 6618-6628. <http://dx.doi.org/10.1016/j.eswa.2010.03.067>
- [5] Gao C, Jian L, Luo S. Modeling of the thermal state change of blast furnace hearth with support vector machines. *Industrial Electronics. IEEE Transactions on*. 2012; 59(2): 1134-1145. <http://dx.doi.org/10.1109/TIE.2011.2159693>
- [6] N. Chauhan, V. Ravi, D. Karthik Chandra. Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks. *Expert Systems with Applications*. 2009; 36: 7659-7665. <http://dx.doi.org/10.1016/j.eswa.2008.09.019>
- [7] T. Lei, R. Guo, X. Wang et al. Application of grey system theory in agriculture, Henan Science and Technology Press.
- [8] Z. FENG and J. Wang. Development of the scale of patients' satisfaction with nursing care using the Delphi method (in Chinese). *Journal of Nursing Administration*. 2007; 7: 6-9.
- [9] T. L. Saaty. 1990: How to make a decision: The Analytic Hierarchy Process. *European Journal of Operational*.
- [10] W. Han and J. Li. Selection of forecasting alarm index for financial distress based on entropy method(in Chinese). *Journal of Beijing Jiaotong University Social Sciences Edition*. 2006; 6: 65-68.
- [11] M. Lehtokangas. Inverse PCA method for weight initialization in CMLP network *International Joint Conference on Neural Networks*. 1999; 2: 878 - 881
- [12] J. Wu, X. Wu and Z. Gao. Multi-attribute weight allocation based on fuzzy clustering analysis and rough sets. *Advances in Computation and Intelligence, Lecture Notes in Computer Science*. 2009; 5821: 358-365.
- [13] K. H. Hsia, M.Y. Chen, M.C. Chang. Comments on Data Preprocessing for Grey Relational Analysis. *Journal of Grey System*. 2004; 7(1): 15-20
- [14] Y. Chen, H. Ke, Y. Li. Grey relational analysis algorithm on weights in multi-attribute group decision-making. *IEEE International Conference on Grey Systems and Intelligent Services*. 2009: 1029-1032.
- [15] L. Zhang, C. Zhang and T. Shi. A Study on Gray Relational Analysis of Many Factor Weights in Tobacco Leaves Classification. *Advanced Materials Research*. 2010; 139-141: 1728-1731.
- [16] J. Kennedy, R.C. Eberhart. Particle swarm optimization. In: *Proc. IEEE Internat. Conf. On Neural Networks*, Piscataway, NJ, USA, pp.1942-1948.
- [17] Rashedi E, Nezamabadi-Pour H, Saryazdi S. GSA: a gravitational search algorithm[J]. *Information sciences*. 2009; 179(13): 2232-2248. <http://dx.doi.org/10.1016/j.ins.2009.03.004>