A New Sustainability Index System and Evaluation Model Based on the GA-BP Neural Network Model

Yiheng He[1, a],

[1]School of North China Electric Power University, Baoding 071003, China.

[a]619132065@qq.com

Keywords: Sustainable development, GA-BP neural network model.

Abstract. This paper aims to evaluate the sustainability of the country using mathematical methods. First we select nine indexes from natural resources, economic strength, human resources and social living environment in total to establish the index system for sustainable development and validate the completeness and independence among indexes. We randomly select ten countries as the sample data, and build an evaluation model based on the GA-BP Neural Network Model [1] which has the advantage in assigning weights objectively, and define a mathematical standard for sustainability. The evaluation results of this model is similar to the conclusion of World Bank.

Introduction

Sustainable development is defined by the 1987 Brundtland Report as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs." [2]The United Nations (UN) predicts the world's population will level at 9 billion people by 2050. This, coupled with increased consumption, places a significant strain on the earth's finite resources [7].Therefore, we must take the path of sustainable development. Decreasing personal poverty and vulnerability, encouraging economic development, and maintaining ecosystem health are the pillars of sustainable development [7].

Sustainable development is not equivalent to the simple ecological or environmental protection. It generally consists of environmental factors, social factors and economic factors of three components. Therefore, a comprehensive sustainability evaluation system should take into account various indicators, as a comprehensive, scientific reaction of the sustainable development of a country. Therefore, we first select nine indexes of four kinds to establish the index system for sustainable, and the completeness and independence are validated among indexes. Then we need to establish a proper mathematical model with conciseness and rigor to evaluate the sustainability of the country according to the above indexes.

The Sustainability Evaluation System

Indexes. We propose a sustainability evaluation system based on the national wealth [4]. The system takes the natural resources, the current economic strengths, human resources and social living environment of four factors into account, including nine sub-indexes as follows:

Natural resources:

1. Forest coverage (%) (FC): referring to the percentage of a country's land area is forest area accounts, which is a reflection of a country or region's forest area occupancy or the richness of forest resources and the degree of realization of the green indicators.
2. Annual primary energy (Total primary energy production per year quadrillion BTU) (TPEP): The index is the total of national oil, gas and coal production of primary energy converted into energy, represented by the energy units. The storage capacity of it is conducive to long-term stable development of the industry.
3. Renewable internal freshwater resource per capita (m^3) (RIFR)

The current economic strength:

4. Gross domestic product per capita (GDP per capita US$)
5. The prevalence of malnutrition [3] (PM %): In developed countries, the value of index is generally zero, while it is generally higher in LDC countries.

Human resources:

6. School enrollment, tertiary (% gross) (SET): The index reflects the general quality and intellectual level of people in the country, which restrict the development of high value-added high-tech industries.

Living environment:

7. Peace index [3]: The index scores for each country to the extent of peace and stability of life. The score results distribute between 1 and 5. A low score indicates that the country is more "peace and stability".
8. Life expectancy at birth year (LEB year): The higher life expectancy prove the better the country's medical conditions and health conditions.
9. Proportion of population using improved drinking-water sources [4] (PPUIDS %): the index intuitively reflect the health situation in the country.

Above-mentioned nine indexes, malnutrition prevalence and peace index are negatively correlated with the results of the evaluation, we call them cost indexes. Other indexes are positively correlated with the results, which we call efficiency indexes.

Data. We first randomly selected thirty countries some of which miss part of the data. Finally, we chose 10 countries having complete data as samples. All data [3-4] [6] about the indexes is in the Table 1.

Table 1 Statistics of ten countries (2012)

	NR			CES		HR	LE		
	FC/%	RIFR/m^3	TPEP/QBTU	GDP/US$	PM/%	SET/%	LEB/year	PPU/%	PI
Guinea	26.3	19736	0.005	493.5	18.2	10	55	52	2.272
Senegal	43.6	1880	0.004	102.3	13.9	8	63	60	2.061
Madagascar	21.4	15116	0.007	445	31.2	4	64	29	2.124
Nepal	25.4	7214	0.031	699.1	12.1	14	68	66	2.001
Cambodia	55.7	8113	0	945.5	16.9	15	71	22	2.207
Haiti	3.6	1279	0.002	775.5	50	2	62	61	2.179
Island	1.4	530064	0.165	44221.7	0	82	83	100	1.113
Switzerland	31.6	5052	0.63	83295.3	0	56	82	100	1.349
Austria	47.3	6524	0.55	48348.2	0	72	80	100	1.337
Colombia	54.3	47585	4.506	7763	14.9	45	73	88	2.625

From Table 1 we can conclude that value of RIFR of Island is considerably larger than that in other countries. When fresh water resource can meet human need, larger resource has no significant effect on sustainability. If

$$x_{ik} \gg x_{ij} \quad j \neq k ,$$

Then

$$x'_{ij} \to 0, x'_{ik} \to 1 .$$

In this case, the index will be useless. Therefore, we replace $\log x_{ij}$ with x_{ij} .

GA-BP Neural Network Model

Introduction. There are many factors involved in the evaluation of sustainable development. Usually the relationship between evaluation and assessment ranks is nonlinear. Artificial neural networks is a very effective tool to solve the nonlinemar problems. [1] Therefore, we choose BP neural network to address this evaluation issue. But we usually meet some problems inherent such as networks often encounter some slow convergence and easily fall into local minimum when using gradient descent to train the BP network [5]. To overcome these shortcomings, we combined genetic algorithm and BP neural network to establish a GA-BP Neural Network Model. First we pre-train globally BP neural network weights and thresholds with a genetic algorithm. Then we get the final weights and threshold by using BP algorithm.

The BP neural network we applied only has a single hidden layer, because single hidden layer model is able to solve more complex input-output problem. The neurons of hidden layers and output layers receive weighted sum of input values and output values through a transfer function.

The formulation of networks output result：

$$\hat{y}_i = g\left(b_2 + \sum_{j=1}^{M} w_2(j) f\left(\sum_{k=1}^{P} w_1(j,k)x_{ik} + b_1(j)\right)\right) \quad (1)$$

Where P is the number of neurons of input layer, M is the number of neurons of hidden layer, $b_1(j)$ is the threshold of hidden layer, b_2 is the threshold of output layer, $w_1(j,k)$ is the weight linking input layer and hidden layer, $w_2(j)$ is the weight linking output layer and hidden layer [4].

The error of the network E：

$$E = \frac{1}{2}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad (2)$$

Where $\hat{y}_i$ is the expectation value of the output.

Besides, we use real-coded each of the weights and thresholds linked to the composition of chromosomes. Coding respectively correspond to weights of input layer to the hidden layer, weights of hidden layer to the output layer, thresholds of hidden layer and threshold of output layer.

Parameters. Main parameters of the algorithm：

- BP neural network structure design：based on the evaluation index system, input layer nodes is 9, the output layer nodes is 1, the hidden layer nodes n is determined by the formula (3)

$$n = \sqrt{n_i + n_o} + d \quad (3)$$

Where n_i is the number of input nodes, n_o is the number of output nodes, d is a constant range of 0~10.Ultimately, we determine the value of n equals 8, Topology of BP neural network is 9-8-1.

- Parameters of GA：Population scale 200, Max generation 500.

Classification. Because evaluation have different dimension, we first standardize the evaluation index in 4.2 in order to guarantee all the input data points in range [0,1].

Efficiency indexes： $$x'_{ij} = \frac{x_{ij}}{x_{i\max}} \quad (4)$$

Cost indexes： $$x'_{ij} = \frac{x_{i\min}}{x_{ij}} \quad (5)$$

The grade of the sustainability of the country does not have a common evaluation standard. We interpolated linearly interval based on the maximum and minimum of all the samples. Then we set impact levels and generate training data for sustainable development.[1] Initially, we set the level of sustainable development 12 ranks, corresponding model output of 1-12 scores. Number 1-10 is the training sample, 11 and 12 are the test sample, sample input values as shown in Table 2.

Table 2 GA - BP interpolation linearly within the training samples and testing samples

Desired output	Input									Actual output
1	0.66	0.00	0.01	0.54	0.00	0.41	0.00	0.22	0.42	1.02
2	0.69	0.09	0.10	0.58	0.09	0.46	0.09	0.29	0.48	1.99
3	0.72	0.18	0.19	0.63	0.18	0.52	0.18	0.36	0.53	2.98
4	0.75	0.27	0.28	0.67	0.27	0.57	0.27	0.43	0.58	3.98
5	0.79	0.36	0.37	0.71	0.36	0.62	0.36	0.50	0.63	4.99
6	0.82	0.45	0.46	0.75	0.45	0.68	0.45	0.57	0.69	6.01
7	0.85	0.55	0.55	0.79	0.55	0.73	0.55	0.65	0.74	7.02
8	0.88	0.64	0.64	0.83	0.64	0.78	0.64	0.72	0.79	8.02
9	0.91	0.73	0.73	0.88	0.73	0.84	0.73	0.79	0.84	9.01
10	0.94	0.82	0.82	0.92	0.82	0.89	0.82	0.86	0.90	9.98
11	0.97	0.91	0.91	0.96	0.91	0.95	0.91	0.93	0.95	10.92
12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	11.83

11th and 12th sample relative error is apart 0.72% and 1.41%. Because relative error is small, we think the trained GA-BP model could be used to evaluate the level of sustainable development of the country.

For each country, we have nine standardized indicators as inputs to the model, and the model outputs a final score of 0-12 scores. According to the score results, the level of sustainable development of the country is grouped into four levels, the specific classification criteria is shown in Table 3.

Table 3 Classification criteria for sustainable development

	0-3	3-6	6-9	9-12
Grade	I	II	III	IV

The higher the score, the higher the sustainable development. Namely the minimum level of sustainability is I, while the highest level of sustainability is IV.

Results

After establishing the model, we solve the problem by MATLAB. The result of our model is showed in Table 4 .

Table 4 Result of our model

Country	Guinea	Senegal	Madagascar	Nepal	Cambodia
Score	3.578312	3.770568	3.32912	3.92895	4.084474
Grade	II	II	II	II	II
Country	Haiti	Island	Switzerland	Austria	Colombia
Score	3.145684	10.78381	10.24347	11.348	8.418741
Grade	II	IV	IV	IV	III

Model Validation. In order to examine our model, we compare the result of our model with the ranking of 120 countries' national wealth of World Bank [6]. The result of World Bank are indicated in Table 5.

Table 5 National Wealth per capita

Country	Switzerland	Austria	Colombia	Senegal	Haiti	Nepal
Wealth	648241	493080	44660	10167	8235	3802
Rank	1	7	53	88	92	112

We can draw a conclusion intuitively that the result of our model is similar to the result of World Bank. In conclusions, we believe that our model can reasonably evaluate the sustainability of the country effectively.

Summary

Combining data in Table 1 and Table 4,we can draw conclusions as follows：

- There are many common characteristic among Grade II countries, e. g., low GDP per capita, low SET, high PM, low PPUIDS, high PI, lack of natural resources and short LEB.
- There are also many common characteristics among Grade IV countries, and they are opposite to Grade II countries'.
- Comparing with three Grade IV countries, the GDP per capita of Colombia is moderate (7763 US$). But its rich natural resources provide a strong support for long-term industrial development of the country. A higher proportion of university entrance is conducive to the development of high-tech, high value-added industries. Based on above-mentioned analyses, we consider that the sustainability of Colombia is relatively high, which is consistent with the result of our model.

In the case where data is integrated, we can evaluate the sustainability of all countries by our model, and give a list of countries most in need of assistance worldwide.

References

[1] Haiyan Chen, Sijing Cai, Minggui Zheng, Evaluation Of China's Sustainable Development of Energy Based on GA-BP Model [J].ACTA ENERGIAE SOLARIS SINICA, (2010) 09:1220-1224.

[2] World Commission on Environment and Development (1987). Our Common Future. Oxford: Oxford University Press. p. 27. ISBN 019282080X.

[3] Information on http://en.wikipedia.org/

[4] Information on http://www.who.int/

[5] Shoukui Si, Xijing Sun. Mathematical Modeling Algorithms and Applications [M]. National Defence Industry Press, 2013.

[6] Information on http://www.worldbank.org/

[7] Information on http://www.comap.com/