ISSN: 2688-8653

Research on Auxiliary System of Forest Management based on Entropy Weight Method and Multiple Linear Regression

Moyu Zhang[1], Xinyi Zhang[1], Xinmeng Qi[1], Tianxi Wang[1], Yumeng Gao[2]

[1]Panjin Branch Institute of Chemical Engineering, Dalian University of Technology, Panjin, Liaoning, 124000, China

[2]Institute of Economics, Liaoning University, Shenyang, Liaoning, 110000, China

Abstract

Although humans rely on forest resources, what really makes trees disappear is the realization that many land use options bring more value than forests. As time flows, greenhouse gases are continuously released, which causing the Greenhouse Effect. As a carbon sequestration expert, forests make people realize that forests should be our friends worthy of cherishment. We should make reasonable planning for forests. Only if we act shall all be saved. Saving the forest is also saving our future. in order to understand the factors affecting the growth of forest area, we first focus on the data of global forest area and multiple subjective and objective factors affecting forest area. Then, the trend of the data and the weight of each influencing factor are analyzed to establish a multiple linear regression of forest area growth and complete the test. The establishment of the model provides a reliable basis for forest managers to consider the influence of various factors in planning.

Keywords

Carbon Sequestration; Entropy Weight Method; Multiple Linear Regression.

1. Introduction

The greenhouse effect produced into a great threat to life. For example, it will lead to the melting of glaciers, the rise of sea level and the disappearance of some coastal cities. Higher carbon dioxide and warmer temperatures will encourage plants to bloom and produce more pollen, and the allergy season will last longer. We all know it does a great deal of harm to people. To mitigate the effects of climate change, we have to take drastic action to reduce the levels of greenhouse gases in the atmosphere.[1]

In this process, there are two ways to play a major role, one is to reduce greenhouse gas emissions, the other is carbon sequestration.[2] Either way, forests play a vital role. We can sequester carbon dioxide from the atmosphere by making forest products. Although these products have a long or short life, reasonable creation of them may enable us to obtain more benefits. At the global level, appropriate deforestation management strategies will be more conducive to carbon sequestration, while excessive deforestation is not conducive to carbon sequestration. Therefore, it is very important for forest managers to balance the multiple benefits of forest products and the absorption and utilization of greenhouse gases by forests. In this process, many factors need to be considered, such as the species and age of trees, the terrain and environment of the forest, and the benefits and life span of forest products.[3]

2. Influencing Factors Model of Forest Area based on Entropy Weight Method

There are many indicators of forest areas for carbon sequestration and wood supply. According to the hypothesis, each natural indicator has no significant influence on each other and has different influence on the area. Through consulting data and data analysis, the effective forest area is affected by the afforestation area of shelterbelt project, afforestation area of returning farmland to forest project, sand control afforestation area, annual average precipitation, forest deforestation area, forest pest control area and other factors.

The weight of each index is analyzed and compared by means of entropy weight method. The following is the flow algorithm of entropy weight method:[4]

Step 1: Determine whether there is a negative number in the input matrix, and if so, standardize to a non-negative range.

Firstly, let's say that the matrix that we've normalized is X.

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,1} & \cdots & x_{n,m} \end{pmatrix} \tag{1}$$

We have n samples and m indices. If there is a negative value, another method is used to convert it to the non-negative interval, for example, the standardized formula is:

$$\boldsymbol{z}_{i,j} = \frac{x_{i,j} - \mathrm{m}\ \left(x_{1,j}, x_{2,j}, \cdots, x_{m,j}\right)}{\mathrm{m}\ \left(x_{1,j}, x_{2,j}, \cdots, x_{m,j}\right) - \mathrm{m}\ \left(x_{1,j}, x_{2,j}, \cdots, x_{m,j}\right)} \tag{2}$$

Step 2: Calculate the proportion of the ith sample of the JTH index, and regard it as the probability used in relative entropy calculation.

Let's say that the normalized matrix that we got in the first step is:

$$Z = \begin{pmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,m} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n,1} & z_{n,1} & \cdots & z_{n,m} \end{pmatrix} \tag{3}$$

So we compute the probability matrix P:

$$P_{i,j} = \frac{z_{i,j}}{\sum_{i=1}^{n} z_{i,j}} \tag{4}$$

Step 3: calculate the information entropy of each indicator, calculate the information utility value, and obtain the entropy weight of each indicator by normalization.

For the j th indicator, its information entropy is calculated by formula: $e_j = -\frac{1}{\ln n}\sum_{i=1}^{n} p_{i,j} \ln p_{i,j}$. The definition of information utility is: $d_j = 1 - e_j$. Therefore, the greater the information utility value, the more information it corresponds to. The entropy weight of each index can be obtained by normalizing the information utility value:

$$\omega_j = \frac{d_j}{\sum d_j} \tag{5}$$

According to the meaning of the question, we use entropy weight method to calculate the weight of each indicator, and the results are shown in the figure below:

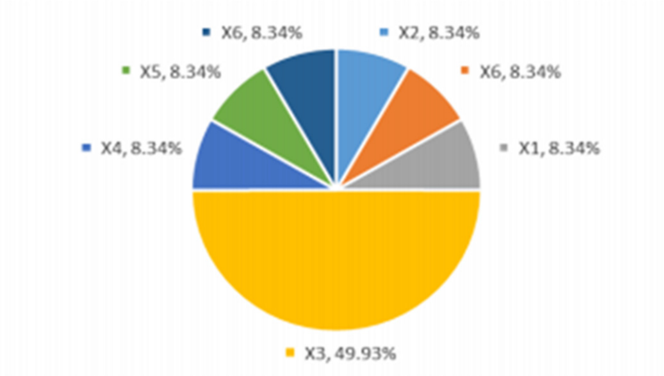

Figure 1. Weight to the Pie Chart

3. Forest Area Growth Model based on Multiple Linear Regression

Multiple linear regression analysis was carried out on seven indexes: shelterbelt forest project afforestation area, returning farmland to forest project afforestation area, sandstorm control afforestation area, average annual precipitation, forest deforestation area and forest pest control area.[5]

Coordinate transformation matrix is the operation of translation, rotation, symmetry, scaling or tangent transformation of a series of points located in different coordinate systems, and finally make these points unified in a plane coordinate system. Here are some basic transformation matrices. Since there is a causal relationship between the driving factors of carbon emissions from energy structure and industrial structure and the resulting factors of carbon emissions from energy consumption, multiple linear regression model is selected after quantitative data processing to establish the model.

Establishment of multiple linear regression model, the multiple linear regression model involving n independent variables can be expressed as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \left(\varepsilon \sim N(0, \sigma^2)\right) \tag{6}$$

Where, β_i is the unknown regression coefficient, and ε is the irregular but normal distribution disturbance term.

If there is a high correlation between explanatory variables in the linear regression model, the model estimate will be distorted or difficult to estimate accurately. If a multiple regression model is to make accurate and efficient estimates, the effects of multicollinearity must be eliminated so that the explanatory variables are not correlated or the correlations are too small to affect the results. In order to eliminate the influence of multicollinearity on regression results and estimate the regression model accurately and effectively, we used SPSS software for multiple linear regression processing. The flow chart is as follows.

Forest in east Asia area as the research object, using the selected seven regressors (a river basin shelter forest project afforestation area, returning farmland to forest project afforestation area, control sand afforestation area, an average annual rainfall, cutting area, forest area of forest

ISSN: 2688-8653

pest prevention and control of diseases and pests), calculated by entropy weight method through Matlab code after each index weight.

Then we use SPSS software to process and get the multiple linear regression equation of forest and related indexes in East Asia.

Table 1. Multiple Linear Regression Equation

Model	Unstandardized Coefficients		Standardized Coefficients	SIG	t	Collinearity Statistics	
	B	*Standard Error*	*Trial Version*			*Tolerance*	*VIF*
CONST	11.958	.000		.	.		
x_1	.003	.000	.303	.	.	.127	7.875
x_2	-5.040E-08	.000	-.082	.	.	.344	2.909
x_3	6.602E-07	.000	1.097	.	.	.086	11.613
x_4	-.005	.000	-2.203	.	.	.036	27.890
x_5	-2.379E-07	.000	-1.156	.	.	.123	8.132
x_6	-.008	.000	-.699	.	.	.442	2.261
x_7	.008	.000	1.923	.	.	.182	5.497

The regression equation is

$$y = 0.303x_1 - 0.082x_2 + 1.097x_3 - 2.203x_4 - 1.156x_5 - 0.699x_6 + 1.923x_7 \tag{7}$$

Where, $x_1, x_2, x_3, x_4, x_5, x_6, x_7$,respectively represent the afforestation area of shelterbelt project, afforestation area of returning farmland to forest project, sand control afforestation area, average annual precipitation, forest deforestation area, and forest pest control area of a certain watershed. SIG is the result of t-test. According to SPSS calculation, the linear relationship between indicators and dependent variables is significant.

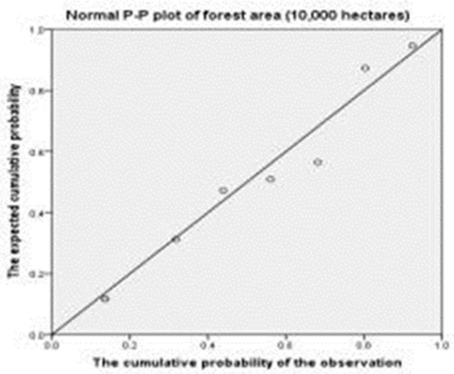

Figure 2. Standard P-P Diagram of Regression Standardized Residuals

From Figure 2, the points on the P-P plot are more concentrated and on the line, so it meets the requirements of regression.

4. Conclusion

According to the model, if we want to restore the forest area occupied by residents for agriculture and industry, or through the construction of shelterbelts and other ecological restoration projects will be destroyed to restore the forest ecological environment is extremely difficult. At the same time, the value of deforestation area is large in the weight analysis, that is,

ISSN: 2688-8653

the current deforestation plan has a great impact on the forest area, so we need to make a correct deforestation and planting plan before the forest ecology is destroyed. The following model was developed to obtain the most satisfactory required forest area, as far as practicable, and to provide cutting and planting options.

References

[1] Li Nan, Wang Jingna, Feng Hongzhi. Reference and enlightenment of Australian energy efficiency and greenhouse effect performance audit [J]. Inner Mongolia Science and Technology and Economy, 2021(10):40-42.

[2] Thomas A. Clark. Carbon capture potential of montane forests: The resilience of local landscapes from a global perspective [J]. Landscape Architecture, 2021, 28(07): 54-68.

[3] Ge Zhipeng. Spatial and temporal variation characteristics and driving mechanism of forest carbon sinks in Zhejiang Province [D]. Zhejiang A&F University, 2021.

[4] Wang Wenqin, Ji Kaiting, Zhang Ye. Regional Land Use Ecological Security Evaluation Based on Entropy Right Method: Taking Chizhou City, Anhui Province as an Example [J]. Journal of Shandong Agricultural Engineering Institute, 2021, 38(06): 22-31.

[5] Zhang Xinhua, Peng Jianhong. Application of Multiple Linear Regression Method in Demand Prediction of Environmental Protection Talents in Poyang Lake Ecological Economic Zone [J]. Shanxi Agricultural Economics, 2016(16):37-38+41.