

# Tree Species Suitability Assessment based on Principle of Maximum Entropy

Wei Xu <sup>a</sup>, Zengchuan Dong <sup>b\*</sup>, Jie Ren <sup>c</sup>, Biao Sun <sup>d</sup>, Yue Zhou <sup>e</sup>, Yiming Wei <sup>f</sup>,  
Fangyu Zhang <sup>g</sup>, Jianting Wang <sup>h</sup>

College of Hydrology and Water Resources, Hohai University, Nanjing, China

<sup>a</sup>hhuxuwei@163.com, <sup>b</sup>zcdong@hhu.edu.cn, <sup>c</sup>jessieren2016@163.com, <sup>d</sup>1498418347@qq.com,  
<sup>e</sup>18351936682@163.com, <sup>f</sup>leonwui@hhu.edu.cn, <sup>g</sup>435103275@qq.com, <sup>h</sup>wjting@hhu.edu.cn

**Keywords:** tree species; suitability assessment; POME; shelter forest

**Abstract.** Factors that affect tree species suitability assessment especially for embankment shelter forest are very complex. And the information for assessment always has incompleteness, multiplicity and uncertainty. To solve the complexity and difficulty of shelter forest vegetation species selection effectively, a tree species suitability assessment model for shelter forest planning based on principle of maximum entropy (POME) using the tree species suitability multi-index system is proposed. And through the proposed model, the tree species can be classified into 4 grades: most suitable, suitable, general and unsuitable. The case study in Baidajie Dike shows that the proposed model is capable of recommending the most suitable tree species for shelter forest and could also provide useful references for the vegetation species selection in landscape planning and design.

## 1 Introduction

Tree species suitability assessment is a hot spot and difficult point in ecological embankment engineering and landscape planning for the assessment information has incompleteness, multiplicity and uncertainty. Currently, the assessment approaches can be divided into 2 categories: qualitative and quantitative. In the qualitative approach, the empirical analysis and the expert recommendation method are now widely used. However, they are all just based on the experts or decision makers' experience. And the reason why a tree species is suitable or not suitable is ambiguous. The uncertainty factors are not fully considered and there is no unified standard for the suitability. In quantitative approach, planting experiment method is the most famous one. And many botanists and foresters have done a lot of experiments to study the tree species' growth habits. But for the ecological embankment engineering design, it may not have a great deal of manpower, materials and financial resources to do the experiments [1], especially in medium and small basins. So this method is not suitable.

Considering the advantages and disadvantages of the two approaches above, many scholars have proposed the approach of combining qualitative and quantitative analysis. And the methods such as fuzzy comprehensive evaluation method, reliability analytical method, artificial neural network method and grey clustering method are proposed and applied in this approach. Among which, the fuzzy comprehensive evaluation method is the most commonly used one, and it has been successfully applied in many other engineering fields [2]. However, the selection of membership functions in this method is subjective, and different scholars may have different judgment criteria, so the membership of each element may be different, resulting different evaluation results. Meanwhile, a lot of useful information may be lost in the maxi-min algorithm of fuzzy set [3].

Entropy, which is originally a thermodynamics concept, was first introduced in information theory by C. E. Shannon [4]. And he defined the average information of the signal in the communication process as entropy to represent the degree of stability, uncertainty and amount of the system information. Principle of Maximum Entropy (POME) was put forward by E. T. Jaynes in 1957 [5]. And in his theory, among all feasible solutions in the ill posed problem, the one which has the maximum value should be selected. Because the maximum entropy means that for the unknown

information, the solution is the most objective and unbiased one under the limited information condition.

Embankment shelter forest is usually planted on the bench land at the water side. The uncertainties in the tree species suitability assessment for the embankment shelter forest are mainly manifested in the following aspects. First, uncertainty caused by the limitations of cognition. Many disciplines and fields are involved in the tree species suitability assessment, such as plant physiology, landscape science, hydraulics, and forest economics. And in these fields, the understandings of the tree species suitability are not the same. Second, uncertainty caused by the insufficiency and unreliability of the collected index value. For some reasons, the index values are insufficient, and sometimes they have to be obtained by reasonable estimation. Third, uncertainty caused by randomness. Both the determination of index system and the obtaining of the index values have randomness. Therefore, POME is an effective way to deal with the problem with these uncertainties.

In this paper, the approach of combining qualitative and quantitative analysis is adopted. First, a tree species suitability multi-index system is established based on the empirical analysis. Then, considering the fuzziness, randomness and incompleteness of the collected information, a comprehensive assessment model that combines the POME and combination weighting method is built and applied in Baidajie Dike. The structure is organized as follows: Section 2 presents the tree species suitability multi-index system. Section 3 presents assessment model based on POME using combination weighting method. Section 4 performs case study of embankment shelter forest tree species suitability assessment in Baidajie Dike, Heilongjiang Province. Finally, Section 5 concludes the paper.

## **2 Tree Species Suitability Multi-index System**

Factors that affect tree species suitability assessment for embankment shelter forest are very complex. And many indexes have been put forward in related research such as landscape architecture, shelter-forest project in north, northeast, and northwest China [6~8].

Embankment shelter forest is an effective barrier to help the dikes against the wind and waves. So the indexes that can reflect wave dissipation characteristics are included as following, and they are (a) Length of main roots (*C1*), (b) Diameter at breast height (*C2*), (c) Crown diameter (*C3*), (d) Flexibility of the branches (*C4*). In order to guarantee the forest survival rate and play a better role in protection function in a certain area, the indexes that reflect growth habit are included as follows, they are (a) Survival rate in flooding for 15 days (*C5*), (b) Growth speed (*C6*), (c) Kinds of diseases and insect pests (*C7*), (d) Cold endurance (*C8*), (e) Soil fertility requirement index (*C9*). Besides, to exert maximal economic and ecological benefits of the shelter forest, Landscape value index (*C10*) and Types of economic value (*C11*) are the representative indexes of other development and utilization values subsystem. Among these indexes, *C7* and *C9* belong to the cost type and the other 9 indexes belong to the benefit type.

Thus the tree species suitability multi-index system was built and its standard was given in Table 1. The index system was classified into 3 groups: wave dissipation characteristics subsystem, growth habit subsystem and other development and utilization values subsystem. And the standard was divided into four grades, and they were (a) most suitable, (b) suitable, (c) general, and (d) unsuitable.

## **3 Tree Species Suitability Assessment Model based on POME**

Suppose there are  $n$  tree species to be evaluated and classified, and each tree species has  $m$  indexes. Then the steps of the classical FOSM are as follows:

- (a) Suitability standard values matrix with  $k$  grades is as follows:

Table 1 Tree species suitability multi-index system

Subsystems	Items	Weights	Most suitable	Suitable	General	Unsuitable
wave dissipation characteristics subsystem	Length of main roots (C1) /m	0.1	1.5	1	0.8	0.6
	Diameter at breast height (C2) /m	0.08	0.2	0.15	0.1	0.05
	Crown diameter (C3) /m	0.1	4.5	3	2	1
	Flexibility of the branches (C4)	0.09	80	60	50	40
growth habit subsystem	Survival rate in flooding for 15 days (C5) /%	0.1	95	85	75	60
	Growth speed (C6) /(m/a)	0.06	1.5	1	0.75	0.5
	Kinds of diseases and insect pests (C7)	0.1	1	3	6	9
	Cold endurance (C8)	0.1	90	70	50	40
	Soil fertility requirement index (C9)	0.1	20	40	50	60
other development and utilization values subsystem	Landscape value index (C10)	0.07	90	70	50	40
	Types of economic value index (C11)	0.1	5	3	2	1

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ s_{21} & s_{22} & \cdots & s_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m2} & \cdots & s_{mk} \end{bmatrix} \quad (1)$$

Convert the suitability standard values matrix above into the fuzzy matrix by formula (2).

$$e_{ih} = \frac{s_{ik} - s_{ih}}{s_{ik} - s_{i1}} \quad (2)$$

Then the fuzzy matrix of the suitability standard values can be got:

$$E = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1k} \\ e_{21} & e_{22} & \cdots & e_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ e_{m1} & e_{m2} & \cdots & e_{mk} \end{bmatrix} \quad (3)$$

(b) Index values matrix of  $n$  tree species is as follows:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} \quad (4)$$

Convert the index values matrix above into the fuzzy matrix by formula (5) and (6) as follows.

For the indexes belong to the cost type:

$$f_{ij} = \begin{cases} 1 & c_{ij} \leq s_{i1} \\ \frac{s_{ik} - c_{ij}}{s_{ik} - s_{i1}} & s_{i1} < c_{ij} < s_{ik} \\ 0 & c_{ij} \geq s_{ik} \end{cases} \quad (5)$$

For the indexes belong to the benefit type:

$$f_{ij} = \begin{cases} 0 & c_{ij} \leq s_{ik} \\ \frac{c_{ij} - s_{ik}}{s_{i1} - s_{ik}} & s_{ik} < c_{ij} < s_{i1} \\ 1 & c_{ij} \geq s_{i1} \end{cases} \quad (6)$$

Then the fuzzy matrix of the index values can be got:

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{bmatrix} \quad (7)$$

(c) Construct the membership degree fuzzy matrix between (3) and (7) as follows:

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{k1} & u_{k2} & \cdots & u_{kn} \end{bmatrix} \quad (8)$$

$$\text{s.t. } \sum_{h=1}^k u_{hj} = 1, \quad u_{hj} \geq 0 \quad (j = 1, 2, \dots, n; h = 1, 2, \dots, k)$$

Obviously, there are infinitely many fuzzy hierarchical matrices that satisfied the constraint conditions. The objective of conducting tree species suitability assessment is that to determine the optimal fuzzy hierarchical matrix. According to the analysis of the uncertainty in the tree species suitability assessment for the embankment shelter forest, the final determination process of  $U$  also has uncertainty. And this uncertainty can be described as Shannon information entropy as follows:

$$H_j = -\sum_{h=1}^k u_{hj} \ln u_{hj} \quad (9)$$

Then the difference between the  $j_{th}$  tree species and the  $h_{th}$  suitability standard can be described as generalized weighted distance.

$$d[e_h, f_j] = u_{hj} \left[ \sum_{i=1}^m (w_i |e_{ih} - f_{ij}|) \right] \quad i = 1, 2, \dots, m \quad (10)$$

(d) Calculate the weight of each index by combination weighting method [9], then the weight vector can be got as follows:

$$W = (w_1, w_2, \dots, w_m)^T \quad (11)$$

The combination weighting method combines the subjective weighting method (the analytic hierarchy process method, AHP) and the objective weighting method (the entropy method, EM) based on game theory. And it avoids not only the subjectivity of the subjective weighting method, but also the absolute objectivity of the objective weighting method.

(e) To determine the optimal fuzzy hierarchical matrix, the generalized weighted distance between the tree species and the suitability standard should be least. Meanwhile, to reduce uncertainty, the Shannon entropy should be maximal. Therefore, the problem of optimal classification is essentially a multi-objective problem. And the optimization model can be expressed as follows:

$$\min_{u_{hj}} \left\{ \sum_{j=1}^n \sum_{h=1}^k u_{hj} \left[ \sum_{i=1}^m w_i |e_{ih} - f_{ij}| \right] + \frac{1}{B} \sum_{j=1}^n \sum_{h=1}^k u_{hj} \ln u_{hj} \right\} \quad (12)$$

$$\text{s.t. } \sum_{h=1}^k u_{hj} = 1, \quad u_{hj} \geq 0 \quad (j = 1, 2, \dots, n; h = 1, 2, \dots, k)$$

Among which, the parameter  $B$  is used to balance the two objectives and is determined according to the actual problems.

Then the Lagrange function of the problem can be constructed as follows:

$$L(u_{hj}, \lambda) = \sum_{j=1}^n \sum_{h=1}^k \left\{ u_{hj} \left[ \sum_{i=1}^m (w_i |e_{ih} - f_{ij}|) \right] + \frac{1}{B} u_{hj} \ln u_{hj} \right\} + \lambda \left[ \sum_{h=1}^k u_{hj} - 1 \right] \quad (13)$$

Among which,  $\lambda$  is the Lagrange multiplier. Then the best  $u_{hj}$  can be solved as follows:

$$u_{hj} = \frac{\exp \left[ -B \sum_{i=1}^m (w_i |e_{ih} - f_{ij}|) \right]}{\sum_{h=1}^k \exp \left[ -B \sum_{i=1}^m (w_i |e_{ih} - f_{ij}|) \right]} \quad (14)$$

#### 4 Case Study

Baidajie Dike, located in Tailai County, Heilongjiang Province, is an important part of Nenjiang embankment project. At present, the dike is 6.7 kilometers, and it is expanded based on the original sand embankment, of which the soil material is too loose. So the waves pose a threat to the dike in flood season. Thus, to reduce the wave disasters, it seems particularly important to plan and build embankment shelter forest.

According to the climate and soil characteristics of the study area, 9 common tree species that include Willow (T1), Pterocarya stenoptera C.DC. (T2), Poplar (T3), Metasequoia (T4), Taxodium ascendens (T5), Taxodium distichum (T6), Koelreuteria integrifoliola (T7), Bischofia polycarpa (T8) and Eucommia ulmoides (T9) were considered, and their index values were collected and given in Table 2.

Besides, the final weights of each index in which determined by combination weighting method were also given in Table 2.

Table 2 Index values of 9 tree species and calculation results of weights

Items	Weights	T1	T2	T3	T4	T5	T6	T7	T8	T9
C1	0.1	8.86	5	4.5	5.83	5	2.62	5	4	1.35
C2	0.08	0.21	0.15	0.18	0.19	0.1	0.12	0.15	0.12	0.1
C3	0.1	5	4	3.4	3.5	2	2.6	5	3	3
C4	0.09	90	50	70	50	50	55	70	55	70
C5	0.1	98.9	87.5	96	93.2	92.6	93.3	60	95	80
C6	0.06	1.2	1.8	2	1	1	0.8	0.6	1.5	0.3
C7	0.1	3	8	6	4	5	1	3	6	4
C8	0.1	90	88	90	38	60	95	68	10	92
C9	0.1	10	30	30	30	50	10	50	10	55
C10	0.07	95	90	88	85	80	80	90	85	50
C11	0.1	4	5	5	3	1	2	3	4	3

Then, set  $B = 10$ . The evaluation results of the 9 tree species by using POME were shown in Table 3. From Table 3, we can find that the Willow tree, Pterocarya stenoptera C.DC. and Poplar are the most suitable tree species for shelter forest construction in Baidajie Dike. So they can be recommended in this region. Besides, Metasequoia, Taxodium distichum, Koelreuteria integrifoliola, Bischofia polycarpa and Eucommia ulmoides belong to the suitable grade; Taxodium ascendens belongs to general grade.

Table 3 Evaluation results of 9 tree species suitability

Tree species	T1	T2	T3	T4	T5	T6	T7	T8	T9
Most suitable	0.943	0.686	0.853	0.227	0.032	0.433	0.188	0.405	0.064
Suitable	0.054	0.255	0.132	0.620	0.352	0.456	0.649	0.411	0.574
General	0.003	0.055	0.014	0.140	0.560	0.104	0.139	0.168	0.327
Unsuitable	0.000	0.004	0.001	0.013	0.056	0.006	0.024	0.015	0.035
Grade	Most suitable	Most suitable	Most suitable	Suitable	General	Suitable	Suitable	Suitable	Suitable

#### 5 Conclusions

In this paper, the approach of combining qualitative and quantitative analysis is adopted. First, a tree species suitability multi-index system is established based on the empirical analysis. Then,

considering the fuzziness, randomness and incompleteness of the collected information, a comprehensive assessment model that combines the POME and combination weighting method is built. Finally, we validate the effectiveness of the proposed model through the case study in Baidajie Dike. The model could provide a useful tool for decision-makers in aspects such as ecological embankment engineering design, garden planning and landscape design.

### Acknowledgment

This work is supported by the Applied Technology Research and Development Program of Heilongjiang Province under Grant No.GZ16B031 and No.GZ16B035.

### References

- [1] W. Xu, Z.C. Dong, J.T. Wang, et al, "An improved fuzzy optimum selection model for choosing wavebreak forest species and its application," in IEEE, Information Technology and Mechatronics Engineering Conference, 2017, pp.122-126.
- [2] T.Y. Li, Z.C. Dong, and R. Han, "Ecological health assessment of Heilongjiang River basin based on fuzzy matter element method," Yangtze River, vol. 48, issue 23, pp. 40-44, 2017.
- [3] H. Ruan, Y.H. Zhang, Z.Q. Zhu, et al, "An improved fuzzy method for evaluating stability of highway slopes," Rock and Soil Mechanics, issue 11, pp. 3337-3344, 2015.
- [4] C.E. Shannon, "A mathematical theory of communication," The Quarterly Review of Biology, vol. 27, pp. 379-423, 1951.
- [5] E.T. Jaynes, "Information theory and statistical mechanics I ," Phys. Rev., vol.106, pp. 620-630,1957.
- [6] X.R. Wang, J.H. Fan, and X.S. Wang, "Distribution of main tree species and its relation to water and heat conditions in shelter-forest districts of 'San Bei'," Journal of Ecology, issue 1, pp. 13-17, 1986.
- [7] M.F. Zhao, "On the role of shrub species in the establishment of three north protective forest,"Journal of Desert Research, vol. 13, issue 3, pp.53-57, 1993.
- [8] B.T. Yu, M.C. Qi, "Species selection for ecological garden trees in winter cities," Journal of Zhejiang A&F University, vol. 32, issue 5, pp. 743-748, 2015.
- [9] C.J. Shan, Z.C. Dong, K.M. Fan, et al, "Application of combination weighting method to weight calculation in river health evaluation," Journal of Hohai University(Natural Sciences), vol. 40, issue 6, pp. 622-628, 2012.