Taking Dandelion as an Example: Exploring a Comprehensive Prediction and Evaluation Model for Invasive Species

Zixin Wang[1], Shaoyu Li[2], Yuqiao Song[3], Jinhao Yao[4]

[1]Yangzhou High School of Jiangsu Province,Yangzhou,225002,China

[2]Ningbo Binhai International Cooperative School,Ningbo,315829,China

[3] Tsinglan School,Dongguan,523000,China

[4]Yali Middle School,Changsha,410027,China

Abstract. With the rapid growth of global trade and tourism, the phenomenon of species invasion is becoming increasingly serious. Invasive species often have a significant impact on native ecosystems, disrupting ecological balance and threatening biodiversity. Therefore, it is essential to conduct in-depth research on invasive species. Dandelion, due to its strong reproductive ability, is considered a common invasive plant, but its application value in medicine, food, and ornamental uses cannot be ignored. Therefore, this article aims to establish a mathematical model to explore the relationship between invasive species and human activities and take Dandelion as an example to establish a comprehensive prediction and evaluation model for detailed analysis.

Keywords: Invasive Plants; Differential Equation; SI epidemic model; AHP; TOPSIS.

1. Introduction

Dandelion is a common plant that propagates through seeds dispersed by wind and can quickly occupy open spaces.Dandelion's adaptation to different environmental conditions and abundant seed production make it invasive.Invasive alien species have significant negative impacts on local ecosystems, potentially leading to the extinction of native species, exacerbating the loss of biodiversity, and even posing a public health risk. Therefore, in order to determine the ecological role of dandelion, we need to predict its growth cycle and analyze relevant influencing factors to determine whether it is a species that plays a positive or negative role in the ecosystem.

2. Notations

The key mathematical notations used in this paper are listed in Table 1.

Table 1: Notations Used in this Paper

Symbol	Description
u	the horizontal wind velocity averaged between the top and the ground
$\bar{u}$	the mean of the distribution of u values during the dispersal season
u_z	an instantaneous horizontal wind velocity at some height
u_g	the geometric mean wind speed of a lognormal distribution of u during the dispersal season
σ^2	the variance in wind speeds in the downwind
$\bar{F}$	a constant descent velocity
Q	The total seed population released by a plant during the dispersal season
$[aij]$	the judgment matrix of the overall seven indexes
C	Comparison matrix

CI	Consistency index
RI	Random index
CR	Consistency ratio

3. Wind dispersal model of dandelion

For simplicity, we only consider the diffusion and growth of a single dandelion fruit and 20000 live seeds within an area of one hectare (hereinafter referred to as the "experimental field"), and construct a wind diffusion model. Based on the simple ballistic formula of Equation 1, we represent the total seed yield of a single plant during its propagation season as the variable Q.

$$x = Hu/F \tag{1}$$

Assuming that seeds detach randomly with respect to wind velocity then insertion of Q in the numerator of Eq. 2 defines dQ/du, which is the frequency distribution of detaching seeds in relation to wind speeds. Assume that xF/H is lognormally distributed during the dispersal season, then gives

$$\frac{dQ}{du} = \frac{QH}{x\bar{F}\sigma_u\sqrt{2\pi}} exp\{\left[\frac{\ln(\overline{xF/H\bar{u}_g})}{\sqrt{2}\sigma_u}\right]^2\} \tag{2}$$

The frequency distribution of seeds on the ground can be represented by dQ/dx, and since by Equation 1,

$$\frac{du}{dx} = \frac{\bar{F}}{H} \tag{3}$$

It follows that

$$\frac{dQ}{dx} = \frac{dQ}{du}\frac{du}{dx} = \frac{Q}{x\sigma u\sqrt{2\pi}} \exp\{-\left[\frac{\ln(\overline{xF/H\bar{u}_g})}{\sqrt{2}\sigma_u}\right]^2\} \tag{4}$$

Equation 4 provides the frequency distribution of seeds at different distances from the point source. We will call dQ/dx the dispersal curve. The distribution of trajectories about the median trajectory (arctan $\bar{F}/\bar{u}_g$)will destined to be skewed to the right, and thus the dispersal curve will be similarly skewed.

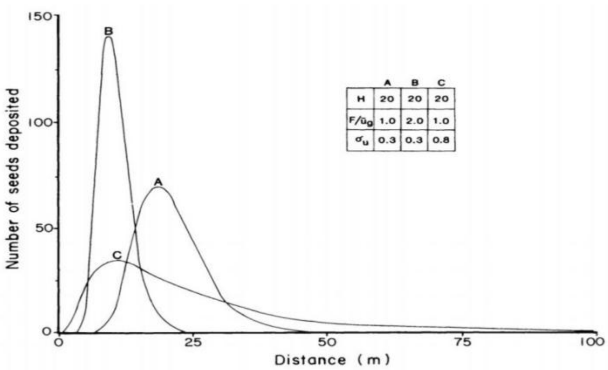

Fig.1: Dispersal curves for differing values of $\boldsymbol{H}$,$\boldsymbol{F/ug}$, and $\boldsymbol{\sigma u}$. As $\bar{\boldsymbol{H}}$ and $\boldsymbol{\sigma u}$ increase, or as $\bar{\boldsymbol{F}}/\overline{\mathbf{ug}}$ decreases, the spread around the median increases.

As is demonstrated in Fig.1 above, if we ignored any effect of σ_w, the relationship between the shape and scale of dispersal curves as a function of $H, F, \bar{u}_g$ and $\bar{\sigma_u}$ shows that as σ_u or H increases, or as $F/\bar{u}_g$ decreases, seeds will be dispersed farther.

According to Walker's model[1], one standard deviation to the right of the mean trajectory is 84%. We have randomly chosen as the distance within which 84% of the dispersed seeds fall ($x_{0.84}$) to characterize the length of the tail. This distance gives an indication of the taper of the cumulative dispersal curve as well as a characterization of long-distance deposition. We can draw the following conclusions from the model algebra. The first result is that a measure of central tendency such as $x_{0.5}$ is a poor guide to the magnitude of $x_{0.84}$.

More than 99.5% of dandelions die within 10 meters for various reasons and are unable to reproduce for the next generation, only 0.05% can fly more than 100 meters, and only 0.014% can fly more than 1 kilometer. As far as the influencing factors are concerned, the strength of the horizontal wind speed has almost no effect on the distance they fly. These results are consistent with Tackenberg[2], as shown in Fig.2.

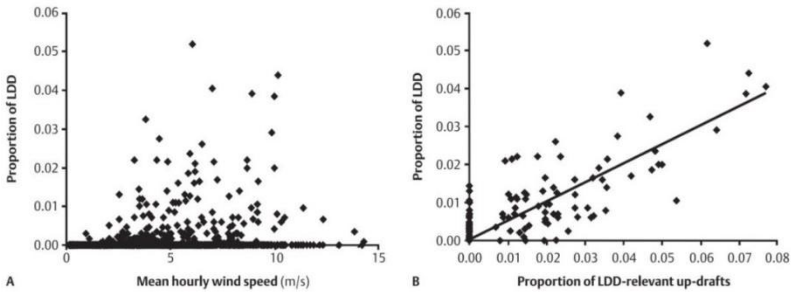

Fig.2: Relationship Between Long-distance Dispersal of Dandelions'Seeds and Wind Speed(A) and Up-drafts(B)

4. SEIDR: Prediction of dandelion spread based on SEIR

4.1 Problem analysis

In order to better analyze the growth of dandelions, we chose to design and implement the model in an open area of one hectare to predict the growth of dandelions. Firstly, we calculated the idealized dandelion growth density according to SHI Y.'s analysis[3], the optimal effective dandelion plant population was 38,352.5 plants/667 m^2 under the 1.5 g/m^2 low-density seeding treatment, i.e., up to 575,000 plants could be present in a hectare. What's more, each individual plant has the capacity to generate a maximum of 20,000 viable seeds.

4.2 Model Establishment

4.2.1 Model structure of SEIR

The SEIR Infectious Disease Model is a common base model used in data modelling, Dandelion seeds are spread in the environment through the wind, mainly considering the path of wind propagation and the mode of seed propagation. Therefore, this section uses an infectious disease model to predict the growth of dandelions. Consider seeds as "sources of infection" and their position in the environment as "individuals". Seeds spread from one place to another through the wind, similar to the transmission of pathogens by infected individuals in infectious diseases.

4.2.2 How the model works

To facilitate the analysis and prediction, some of the parameters used in this model are as follows. The side of the test field was 100 m long and the dandelion plants were 3 cm in diameter. Therefore, the total number of people in the model is 90,000,000.

The "Susceptible" ones are the number of plants that may grow on the land. We idealize the situation that susceptible ones shall be the maximum number of dandelions that can be accommodated per unit area, which is 575,000 here.

"Prevalence" is the percentage of dandelions that spread seed, germinate and grow to the puffball stage on land within one hectare.

The seeds are spread by wind, and the average flight distance of dandelions is 2.15 meters, more than 99.5% of the dandelions' flight distance is within 10 meters, so the “Exposed” is the boundary of the land that is 10 meters away from the previous seed surroundings.

The lifespan of the dandelions is usually two years, but even though the model considers the optimal survival conditions, density-constrained population growth under nutrient- and space-constrained conditions cannot be avoided, so the number of dandelions forced to wither is considered to be the number of the “Recovered”, and in this paper, we set the actual mortality rate at seedling rearing to be 25%.

Virus transmission probability is the probability of dandelion seed transmission, and its value receives the influence of temperature, climate, etc. We will discuss the impact of climate in more detail later and for the time being let the impact factor be X According to the results in subsection 4.2, the rate of transmission varies with distance in three categories, which are 60% in 2.5 meters, another 39.5% in 10 meters and the rest more than 10 meters. Therefore, the average transmission probability was chosen to be $0.6X$,$0.395X$,$0.005X$.

In order to have the best developmental environment at the beginning, we chose to use California as an example and set up the experiment to begin in June, when soil temperatures approach 77°F (25°C), dandelion seeds germinate faster[4], with the germination and seedling stage lasting 8 weeks for the first 4 months, and the setup requiring 12 weeks of germination and growth for the 5th-December months, when temperatures and humidity decrease.

4.3 Results

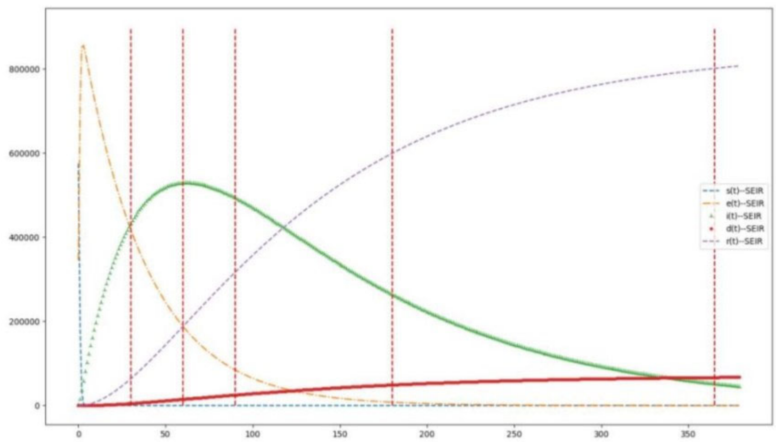

Fig.3: Result of SEIDR Model

As shown in Fig.3 above, the spread of dandelions over the course of 1, 2, 3, 6, and 12 months is predicted as follows. After a month the first dandelion growth and development is complete and about 3,450 new dandelions reach the puffball stage and begin the next stage of wind dispersal propagation, which can result in over 30,000 dandelions in the first three months. As the plants become more densely populated, effective substance accumulation and yield are significantly suppressed, and the rate of dandelion growth and reproduction in this hectare begins to slow down. optimal planting densities are reached in 6 months when the growth rate slows down even further, and there are fewer and fewer open spaces in the experimental plots. Considering the low environmental requirements and high reproductive capacity of dandelions, under the ideal condition of the model design, the hectare will be covered with dandelions after oneyear, reaching about800,000, leading to aplantgap of 11 cm, which isareasonablespacing for growth. If the weather is too cold, with temperatures at -23.4 degrees Fahrenheit, dandelion seeds will go dormant.

5. Top-AHP: Invasive Plant Discriminative Model

As mentioned in the previous section, dandelions grow and reproduce very quickly, and have significant invasiveness in the field of biology. At the same time, dandelion has important medicinal value in the fields of herbal medicine and medicine. Therefore, when exploring the role of dandelion, it is necessary to pay attention to its potential negative impact on the ecosystem and its valuable resource value in the medical field.

5.1 Index Extraction and Calculation

In our model, we constructed a second layer index to measure five aspects of plants to determine whether they are invasive plants. These five aspects include individual characteristics, ecological impacts, and land use characteristics.

5.1.1 Seed Size

The researchers found that the species with small seed mass tend to be more successful invaders.[5] Seed coat thickness, for example, can affect germination rate and imbibition capacity, which are important for successful establishment in new environments.[11] Seed size is another seed trait that can influence invasiveness. Smaller seeds may have advantages in terms of dispersal ability, as they can be easily carried by wind or animals.

5.1.2 Product of Growth Rate and Life Span

We combine the growth rate and life span. The larger the product, the more likely it is that the invasive plant is approximately.

5.1.3 Specific Leaf Area

Specific Leaf Area (SLA) is a measurement used to assess leaf thickness and density and is commonly used to compare different plant species or assessplant responses to environmental conditions. The calculation formula for specific leaf area (SLA) is:

SLA(m^2/g) = Leaf area (m^2) / Dry weight of the plant (g) (5.) Here, leaf area can be obtained by weighing method, image analysis, or other methods,

and the dry weight of the plant can be obtained by weighing method.

High SLA was significantly and uniquely correlated with invasion success at the continental scale.[6] SLA was positively correlated with regional and continental abundance. Since dandelions are terrestrial plants, we used specific leaf area as an important metric. The larger the specific leaf area, the more likely it is to be an invasive plant.

5.1.4 Rarity-weighted species richness index

Rarity-weighted species richness index involves the presence/absence of rare species, which is created by the Kentucky State Nature Preserves Commission. The rarity index score is derived by integrating the distribution patterns of rare species and the count of populations present within the state.

5.1.5 GAP diversity

GAP (Gap Analysis Program) is a comprehensive approach that aims to identify gaps in the conservation of biodiversity by assessing the representation and distribution of habitats and species across a given geographic area. The program utilizes various data sources, including maps, satellite imagery, and species occurrence records, to analyze and model habitat diversity.

The concept of GAP diversity encompasses the characterization and quantification of habitat types within a specific region. The GAP analysis program provides valuable insights into the status and extent of habitat diversity within an area, allowing conservationists and resource managers to identify areas of high conservation value, prioritize conservation efforts, and make informed decisions regarding land use and species protection.

5.1.6 Economic Benefit

This indicator is used to assess the potential impact of dandelions on our economy, synthesizing the market demand for the plant, the size of the industry, and seeing if there is value in areas such as tourism, food, and environmental protection, or if there is a positive or negative impact on daily life.

5.1.7 Construction Cost

Construction cost (CC) is a quantifiable measure of energy demand for biomass production and is related to energy-use efficiency. It has been hypothesized that invaders may gain a growth advantage by efficiently utilizing energy due to their lower construction costs.[8]

5.1.8 Normalization

After acquiring the values of all the indexes, we process the data through normalization. Because our indexes are all benefit indexes, we can use the formula below to process our data directly:

$$a_{ij}^{*} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{n} a_{ij}^{2}}} \tag{5}$$

5.2 Evaluation Model

5.2.1 TOPSIS

TOPSIS, a widely used method for comprehensive evaluation within a group, utilizes complete information from the original data through a sequential preference technique similar to the ideal solution. The results effectively reflect and quantify the gap between evaluation schemes. The basic process of TOPSIS includes utilizing the normalization matrix of the raw data. Apply cosine method to identify the optimal and worst solutions in a finite set of solutions. Subsequently, calculate the distance between each evaluation object and the optimal and worst solutions. This allows for determining the relative closeness between each evaluation object and the optimal solution as the basis for the evaluation process.

Furthermore, the TOPSIS method is not strictly constrained by data distribution or sample content, making it applicable to a wide range of scenarios. Its calculation process is straightforward and user-friendly, facilitating ease of use and implementation.

5.2.2 Analytic Hierarchy Process（AHP）

The Analytic Hierarchy Process, known as AHP, is a model that aids people in making decisions by applying pairwise comparisons.Fundamentally, AHP establishes the priority of alternative solutions and the criteria used to judge them, which helps with the overall decision-making process. Using AHP, the indicators we consider will be arranged in a hierarchical structure.

After calculation, Table 2 lists the importance and corresponding importance levels of a scale from 1 to 9. The scaling method for determining matrix elements is as follows.

The consistency index is defined as:

$$CI = \frac{\lambda - n}{n - 1} \tag{6}$$

Our next step is to calculate the consistency ratio (CR) in order to confirm the results from theAHP. The formula is shown below:

$$\mathrm{CR} = \frac{CI}{\mathrm{R}I} \tag{7}$$

Table 2: Scores for the Importance of Variable

Numerical Scale	The Level of Importance
1	Equal Importance
2	Equal to Moderate Importance
3	Moderate Importance
4	Moderate to Strong Importance
5	Strong Importance
6	Strong to Very Strong Importance
7	Very Strong Importance
8	Very Strong to Extreme Importance
9	Extreme Importance
1/N	The reciprocal of the above values, with meanings relative to the above

To measure the of CI, the random consistency index RI is introduced:

$$RI = \frac{CI\ \ +_1CI_2 + \cdots + CI_n}{n} \tag{8}$$

The value of RI is influenced by the dimension of the matrix, as shown in Table 3

Table 3: The RI Value of AHP

$\boldsymbol{n}$	1	2	3	4	5	6	7	8	9	10	11
$RI\ Value$	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

5.3 Model Results

In this part, we will explain how our model is capable of determining an 'impact factor' for invasive species, and we will show how our model fits the dandelions.

Below is a typical judgment matrix of a botanist whose judgment ranking is **Seed Size > Construction Cost > Product of Growth Rate and Life Span > Rarity-weighted Species Richness Index > Economical Benefit > GAP Diversity > Specific Leaf Area**. The scales we set in the matrices obey the following rules:

1) All the scales are used to compare the degree of importance between two things. For example, if a botanist thinks seed size is better than the GAP diversity, the scale will be: $a_{SG} > 1$.

2) All the scales considering our indexes can be divided into 9 degrees.

3) We all use positive integer scales. When two indexes are compared, the latter might be better than the former, and a fraction occurs. In case the reciprocal of the fraction can also be an integer, the format of the fractional scale should be:

$$\mathrm{aij} = \frac{1}{\mathrm{n}}, \mathrm{n} \in \mathrm{N}^* \tag{9}$$

Thus, the final judgment matrix is shown in Fig.4:

$$[a_{ij}] = \begin{bmatrix} 1 & 3 & 6 & 2 & 7 & 5 & 4 \\ 1/3 & 1 & 4 & 6 & 2 & 3 & 5 \\ 1/6 & 1/4 & 1 & 5 & 3 & 2 & 6 \\ 1/2 & 1/6 & 1/5 & 1 & 4 & 6 & 2 \\ 1/7 & 1/2 & 1/3 & 1/4 & 1 & 2 & 3 \\ 1/5 & 1/3 & 1/2 & 1/6 & 1/2 & 1 & 4 \\ 1/4 & 1/5 & 1/6 & 1/2 & 1/3 & 1/4 & 1 \end{bmatrix}$$

Fig.4 the final judgment matrix

Then we calculate the eigenvector of the matrix, which is [0.261, 0.174, 0.050, 0.158, 0.106, 0.111, 0.195], indicating a larger weighting for the Seed Size index than others. Studies have shown that plants with small seeds tend to have strong dispersal ability, and due to small seeds, a large number of individual plants, and stronger resistance to diseases and pests, a certain number of seeds can still survive and grow even if the environment changes considerably. At the same time, these plants can rapidly expand their population size through asexual and sexual reproduction. These characteristics enable them to rapidly take root in new ecological environments, expand their population size, seize ecological resources, and even have a serious impact on local ecosystems.

The maximum eigenvalue of the above judgment matrix is 7.08, thus leading to the consistency index CI =0.01, corresponding to the consistency ratio CR= 0.0076<0.1. In this case, this level of consistency can be accepted, and this level of consistency is acceptable in the practical application of the judgment matrix. After transforming the characteristics of the plants into the preference variables, judgment matrices, and conformity variables, we can calculate all seven indexes of each plant and score other plants through TOPSIS. Among all the indicators, Seed Size, GAP Diversity and Economic Benefit, these three indicators need to be positively normalized first. After calculating the positive and negative ideal solutions and calculating the distance, we can get the closeness. The closer the resulting value of relative closeness is to 1, the closer the sample plant is to a positive ideal solution and the more likely it is to be recognized as an invasive species.

6. Model Application

6.1 Application for Dandelions

Then we should test the model by using it to compute an impact factor for dandelions. We score Dandelion as the following Table 4.

Table 4: Model Test for Dandelions

	Parameter	Normalized -Score	Description
	Seed Size	85	Given that dandelions have small, lightweight seeds that can be carried by wind and are therefore efficient invaders, they could be scored highly.
Individual Charateristics	Product of Growth Rate and Life Span	70	Dandelions grow rapidly and have a long lifespan as perennials, potentially several years.
	Specific Leaf Area Rarity-weighted	85	Data obtained from references
Ecological Impacts	Species Richness Index	70	Integrating that the dandelion's universal but widespread
	GAP Diversity	15	Dandelions can thrive in diverse habitats. Their ability to grow in various conditions could damage the ecological environment Dandelions can sometimes have economic
Economical Impacts	Economic Benefit	50	importance due to their use in herbal medicine, food and horticulture, but it's usually limited.
	Construction Cost	80	Data obtained from references

Finally, we calculated the “Invasive Index” of Dandelions, whose score is 0.686. As the score is more than 0.500, dandelions should be classified as an invasive species. However, this score is not too large. Therefore, it can be seen that the relationship between dandelions and us humans is both competitive and cooperative, beneficial and harmful. As a plant, dandelion has strong vitality and adaptability, and its seeds can be spread to various places by wind to reproduce and grow. This characteristic makes dandelion highly competitive for survival in the ecological environment.

On the one hand, as a plant, dandelion can absorb carbon dioxide, release oxygen and purify the air, which has a certain effect of improving our ecological environment. Meanwhile, the medicinal value of dandelion is also widely recognized, with the effects of clearing heat and detoxification, anti-inflammatory and pain relief, which can be used as herbal medicine. In addition, dandelion is edible and has certain nutritional value.

On the other hand, the growth of dandelion will also produce competitive pressure on other plants, seizing resources such as sunlight, water and nutrition. Under certain conditions, excessive growth of dandelion may result in limiting the growth of other plants, thus affecting the ecological balance. In addition, the seeds of dandelion are irritating and prone to triggering allergic reactions, which may cause distress to some people.

To summarize, the relationship between dandelion and us human beings is a state of both enemy and friend. While rationally utilizing and protecting dandelion, we also need to pay attention to its possible negative impacts and take measures to balance the various relationships in the ecological environment, so as to realize the harmonious coexistence of human beings and nature.

6.2 Application for Another Two Invasive Plants

To better validate the modelling effects, we apply our model to determine the impact factor for two other plant species, Tumbleweed and Purple loosestrife that are often considered invasive.

We substituted two plants into the model to go for a fuller evaluation, these two plants are scored as Fig.5.

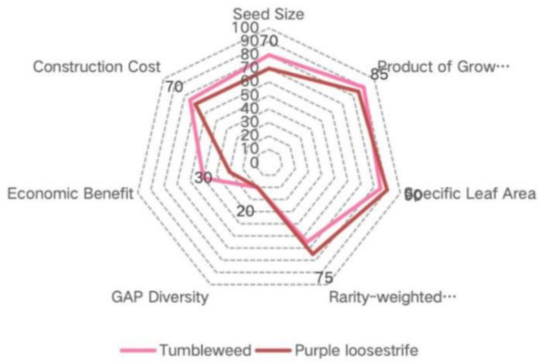

Fig.5: Invasive Index Score for Tumbleweed and Purple loosestrife

The final calculations yielded scores of 0.705 and 0.649 respectively, in other words, Tumbleweed and Purple loosestrife are both considered invasive plant species in certain regions.

7. Conclusion

7.1 Summary of Results

Dandelion, Tumbleweed and Purple Loosestrife are all classified as invasive plants in North America and do have a negative impact on the local ecosystem. The ability of these invasive plants to reproduce rapidly gives them an advantage in competing for sunlight, water, nutrients, etc., thus inhibiting native plant populations and affecting the ecological balance. Therefore, when facing these invasive plants, we should adopt a rational and scientific attitude to prevent their excessive impact on the ecological environment, fully utilize their ecological value and human interests. In the process of dealing with invasive plants, various methods such as biological, chemical, and physical control can be adopted to reduce their reproduction rate and transmission range, and protect local plant populations. Meanwhile, by studying the biological characteristics of these invasive plants, we can find effective control measures, thereby providing scientific basis for global ecological environment protection and plant resource management.

7.2 Strengths

The wind spread prediction model based on distance is scientific and reasonable, and can pass various statistical tests. The predictions obtained have a reliable statistical description;

We use both AHP and TOPSIS to comprehensively consider whether a plant is an invasive plant from both subjective and objective perspectives, and this evaluation method has high accuracy and comprehensiveness, which helps us to better recognize the invasive characteristics of plants and provides a scientific basis for ecological environmental protection, plant resource management and invasive plant control.

The E-AHP model is very convenient for the government or the household to make rational decisions according to the actual situation.

7.3 Possible Improvements

The analysis of prediction and evaluation can be more accurate if we have more complete data and if our qualitative analysis is done by experts;

Some approximate analysis methods are applied to model the evaluation of invasive plants, which may lead to a situation contrary to the actual in extreme cases.

References

[1] Walker, E. R. (1965). A particulate diffusion experiment. Journal of Applied Mete-orol-ogy (1962-1982), 614-621.

[2] Tackenberg, O., Poschlod, P., & Kahmen, S. (2003). Dandelion seed dispersal: the hor-izontal wind speed does not matter for long-distance dispersal-it is updraft!. Plant Biology, 5(05), 451-454.

[3] SHI Y., YANG L., Li L, & et al.(2021). Effects of Sowing Densities on Biomass and Yield of Dandelion under Forest. Gansu Agricultural Science and Technology (10),23-26. doi:CNKI:SUN:GSNK.0.2021-10-007.

[4] Dreistadt, S., Phillips, P., & O'Donnell, C. (2004). Integrated pest management for homegardenersand landscapeprofessionals. Pestsof landscapetreesand shrubs:An integrated pest management guide.

[5] Moles, A. T., Falster, D. S., Leishman, M. R., & Westoby, M. (2004). Small- seeded species produce more seeds per square metre of canopy per year, but not per individual per life-time. Journal of ecology, 92(3), 384-396.

[6] Hamilton, M. A., Murray, B. R., Cadotte, M. W., Hose, G. C., Baker, A. C., Harris, C. J., & Licari, D. (2005). Life-history correlates of plant invasiveness at regional and con-tinen-tal scales. Ecology Letters, 8(10), 1066-1074.

[7] Song, L. Y., Ni, G. Y., Chen, B. M., & Peng, S. L. (2007). Energetic cost of leaf con-struction in the invasive weed Mikania micrantha HBK and its co-occurring species: im-plica-tions for invasiveness. Botanical Studies, 48(3), 331-338