Team Control Number

14079

Problem Chosen

A

2023

HiMCM

Abstract

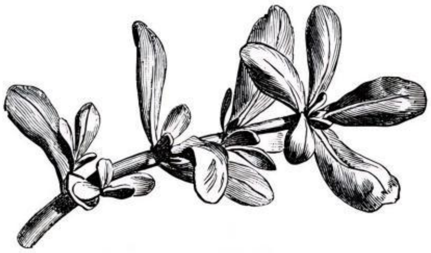
Taraxacum mongolicum, Dandelion, is always known for its **seed propagation**. When the wind blows, seeds contained in "white cotton" fly into the sky and could land anywhere possible, even to locations in which dandelion is **non-native**. Since resources are finite in a habitat; thus, "**prosperity**" of dandelions might deplete most of the water and nutrition which used to belong to **native species**. Then, Dandelion could be considered as "aggressor" or even to a professional term in Biology and Geography: **Invasive Species**.

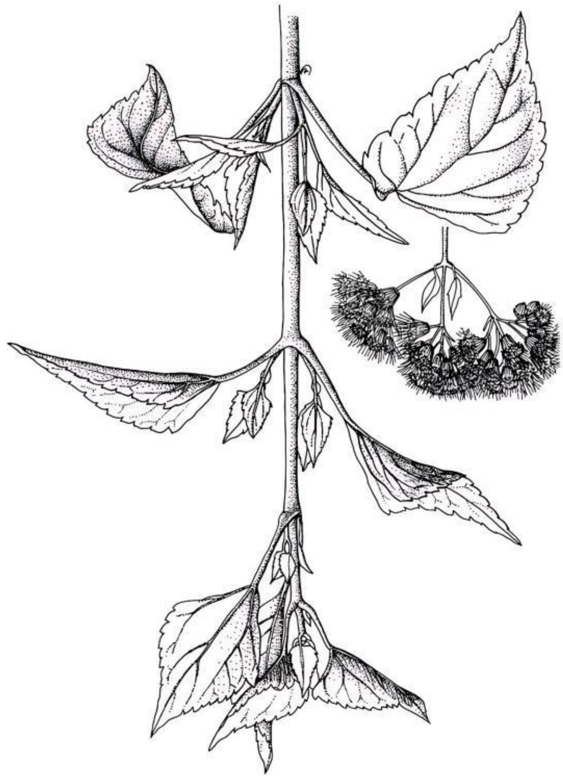
Determining whether Dandelion is invasive or not is controversial since they are not similar to common invasive species who inactivate soil fertility or disrupt ecological balance. They mainly spread far, quickly and easy to grow. So, all around the world, few cities consider them as invasive species.

Nonetheless, Dandelion's "unscrupulous" dispersion might replace all of the flowers in one's garden; thus, they could be **"invasive"** to individuals. In order to help people to identify species similar to dandelion and then prevent tragic replacement in garden, we mainly focus on simulating the Dandelion Dispersion model and develop an App based on that.

We apply Probability **Density Functions** and **Normal distribution** for the first question, mainly focusing on **wind direction** and **wind speed** and considering the effect of all other factors merely affect the **maturity rate**. For question 2, **AHP** and **TOPSIS** are utilized creatively to derive **Threaten Coefficient** of each species and then justify whether they are invasive species or not.

Keywords: Wind Speed, Wind Direction, Offspring numbers, PDFs, ND, AHP, Threaten Coefficient, and TOPSIS.





AHP

TOPSIS

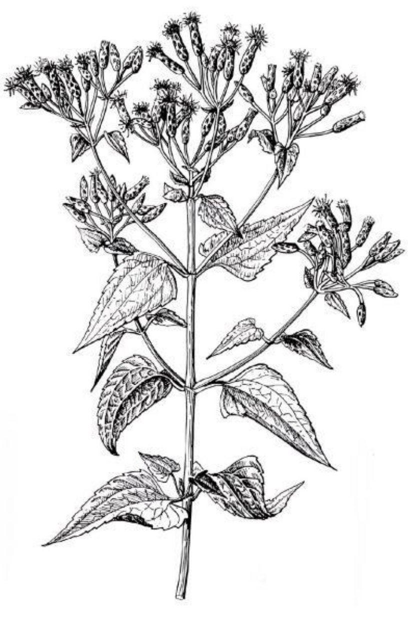
Friends or Foe?

Both or Neither?

Normal Distribution

Probability Density

Functions



#14079 HIMCM

Content

1 Introduction **1**

2 Problem Restatement and Analysis **1**

2.1 Dandelion Dispersion Model (DDM) 1

2.2 Species Invasive Coefficient Evaluation Model(SICEM) 1

3 Assumptions and Justification **1**

4 Dandelion Dispersion Model (DDM) **3**

4.1 Variables 3

4.2 Typical Dandelion Life cycle 3

4.3 Mathematical Modeling of Dispersion 3

4.5 Results 9

4.6 Sensitivity Analysis 10

5 Species Invasive Coefficient Evaluation Model(SICEM) **12**

5.1 Variables and Definitions 12

5.2 TOPSIS 13

5.3. AHP 14

5.4 Results&Conclusion 16

5.5 Sensitivity Analysis 17

5.6 Result 18

6 Application **18**

7 Strengths and Limitations **20**

7.1 Limitations 20

7.2 Strength 20

References **22**

1 Introduction

Transported from Europe and Asia, dandelion could be used as medicine and food. Besides, it is always a meaningful and aesthetic symbol for freedom in literature due to its splendid way of seed dispersion. And because of its strong life capability, it is also considered as the symbol of bitter-ender who never chooses to give up after failure.

On the other hand, when dandelion is delivered by wind to another space, native species here are probably doomed to die as a result of unchangeable limitation of resources in a land and dandelion's strong life ability provided by them long roots. Consequently, whether dandelion is an invasive species requires us to ponder carefully by further and careful investigation.

2 Problem Restatement and Analysis

2.1 Dandelion Dispersion Model (DDM)

We are required to construct a mathematical model (DDM) that estimates the propagation situation over time from a single dandelion in its "puffball" stage adjacent to an open one-hectare plot of land. There are 2 main stages of Dandelion Life Cycle : Growth Stage and Reproduction Stage; thus, we are supposed to take the time interval between those stages into account as fundamental information for our model. We should also consider the main factor affect the dispersion: Scale related to wind direction as well as wind speed and population related to maturity rate.

Conduct sensitivity analysis is done eventually to determine applicable scope of our model. More specifically, it will tell the change of which initial variables lead to the most significant change in the seed dispersion when other factors are stable and normal.

2.2 Species Invasive Coefficient Evaluation Model(SICEM)

We are asked to build a mathematical model(SICEM) applicable for determining an 'impact factor' for invasive species. Various factors, such as harmfulness and spreadability, are supposed to be taken into consideration. Besides, reusing DDM could offer more thorough results. Thus, we should test our SICEM on species often considered invasive.

Based on the data we acquired in the previous questions, we should put together a straightforward blog that individuals could use as a potential guide. In the blog, we should point out the factors that have the most influence over the seed propagation and give our estimate of the approximate population numbers and distribution in an acre of land.

3 Assumptions and Justification

- **Assumption 1**: A plant of dandelion is adjacent to an acre of open wasteland, without dominant predators and human beings.

- **Justification 1**: To simplify the model, we choose to ignore the impact of human behavior on dandelions. Since this unpredictable factor could not be converted into quantitative factor, we choose to ignore this secondary contradiction.
- **Assumption 2**: The initial location of dandelion is randomly distributed on the edge of a hectare of land.
- **Justification 2**: according to the question, dandelion is adjacent to a hectare of land. Random distribution is able to reflect the impact of different locations and wind speeds on the spread of dandelions more intuitively.
- **Assumption 3**: The maturity rate of Dandelion is 2%.
- **Justification 3**: The propagation situation included the scale of dispersion and the number of offspring. We assume that wind speed and wind directions affect the transportation distance while other factors such as natural predators, seed burial depth merely affect the maturity rate. We derive the maturity rate as a constant for model simulation, according to reading in essays.
- **Assumption 4**: It is a plain without apparent fluctuation of height.
- **Justification 4**: A high hill, even a mountain is able to buffer the dispersion of dandelion though wind.
- **Assumption 5**: An uncommon and emergent case does not happen.
- **Justification 5**: Data we derived are usually average value eliminating the outliers, thus, it makes sure our model fits most cases and is universal.
- **Assumption 6**: We do not consider the maximum number of dandelions per square meter
- **Justification 6**: Since the dispersion speed is so fast and the land area is relatively small, setting an upper limit will not show the characteristics of dandelion propagation.
- **Assumption 7**: Set the seed release height H of this dandelion is 0.5 m and the seed sedimentation rate F is 0.5
- **Justification 7**: By reviewing the relevant data of seed release height and sedimentation rate of six dandelion species [7~9], we established a general model. Setting H and F as constants according to [10] simplifies the model.
- **Assumption 8**: A month only has 30 days.
- **Justification 8**: To simplify the calculation, we ignore the difference in the length of each month and unify it to 30 days
- **Assumption 9**: The dandelion breeding season is from January to May
- **Justification 9**: Due to the characteristics of the breeding season of dandelion: mainly in spring, the change after May is not apparent.
- **Assumption 10**: Dandelion reproduce 4 times a year.
- **Justification 10**: Dandelion reproduce more than one times a year. Perennial dandelion can produce more and more times as they grow up; thus, we take the average number.
- **Assumption 11**: The standard deviation of wind speed each month is derived as a random function.
- **Justification 11**: Since changes in wind speed each month are difficult to monitor, we set up a random function to determine the standard deviation of wind speed.

4 Dandelion Dispersion Model (DDM)

4.1 Variables

Variable	Description
Q	Number of seeds carried by a dandelion and the number of dandelion seed wind dispersal trials
$\varphi(x)$	The number of seeds that have dispersed and accumulated at a distance extending from x to $(x+\Delta x)$ meters from the point of origin of the dandelion seed dispersal
$\Phi(x+\triangle x)$	The probability of a random event occurring within the range of x to $\frac{(x+\Delta x)}{\Delta x}$ meters from the seed dispersal source
x	Dispersion distance
μ	Wind speed
F	Seed descent rate
H	Seed release height
η	Random variable of wind-dispersed distance
ξ	Random variable of wind speed
ψ	Natural logarithm of the wind speed random variable ξ
σ_x	Standard deviation of the random variable ψ
μ_x	Mean of the random variable ψ
Δx	the change or difference in the value of x

4.2 Typical Dandelion Life cycle

We divide the life cycle of dandelion into 2 stages: Growth Stage and Reproduction Stage. The interval between those stages are about 28 days, almost a month.

4.3 Mathematical Modeling of Dispersion

1

The distribution of the number and intensity of seeds deposited and accumulated at different distances from the source during the wind-propagation process of plant seeds is called the wind-borne curve of plant seeds[1]

In order to analyze the characteristics of the wind-propagation curve of dandelion seeds, we assume that the number of seeds carried by the initial dandelion is Q grains. Assuming that seeds propagate merely in one direction under the influence of wind, we set the number of seeds that settle and accumulate in the range from x to $(x+\Delta x)$ m from the initial dandelion seed propagation source be $\varphi(x)$ grains, which is the function of the propagation distance, and set the proportion of the number of seeds that spread within this distance range to the total number of seeds $\Phi(x+\Delta x)$, which is:

$$\Phi(x+\Delta x)=\frac{\varphi(x)}{Q} \quad [1\sim3]\ [6]$$

2

Obviously, $\Phi(x+\triangle x)$ is a function with respect to the distance traveled. Since changes of $\varphi(x)$ with wind has apparent randomicity, it can be considered that $\varphi(x)$ is a random variable, and the proportion $\Phi(x+\triangle x)$ is considered to be the frequency of the random event (expressed in $\varphi(x)$) of the number of seeds that have been deposited and accumulated at a distance of x to $(x+\triangle x)$ meters from the seed dispersal source in the repeated test of Q dandelion seed wind propagation.

As the number of trials increases, the frequency $\Phi(x+\triangle x)$ of this random event $\varphi(x)$ will approach a constant, which can be considered as the probability of random event $\varphi(x)$, from x to $(x+\triangle x)$ meters from the seed propagation source, or the probability of seed wind propagation from x to $(x+\triangle x)$ meters from the seed propagation source, according to the definition of frequency stability and the probability of the occurrence of random events. After knowing the probability $\Phi(x+\triangle x)$, we can calculate the number of seeds that have settled and accumulated from x to $(x+\triangle x)$ meters from the seed propagation source according to the total number of seeds Q, which is:

$$\varphi(x) = Q \times \Phi(x+\triangle x)$$

3

If the seed propagation distance x is a continuous random variable η, the probability density is $g(x)$, and the distribution function is $G(x)$, then the probability of random event $\Phi(x)$ occurring in the range of x to $(x+\triangle x)$ meters is:

$$\Phi(x+\triangle x) = \int_x^{x+\Delta x} g(x)dx$$

4

Substitute Equation. (3) into Equation. (2), which is:

$$\varphi(x) = Q \times \int_x^{x+\Delta x} g(x)dx$$

5

Divide both sides of Equation. (4) by $\triangle x$ at the same time, which is:

$$\frac{\varphi(x)}{\Delta x} = \frac{Q \times \int_x^{x+\Delta x} g(x)dx}{\Delta x}$$

6

When $\triangle x \to 0$, find the limit of both sides of Equation. (5) at the same time, then:

$$\frac{\varphi(x)}{dx} = Q \times \frac{(G(x+\Delta x) - G(x))}{dx}$$

7

According to the derivative of the function and the definition of the probability density of the distribution function of random variables, then:

$$\frac{\varphi(x)}{dx} = Q \times g(x)$$

8

Equation. (7) reflects that the ratio of the number of seeds that accumulate due to wind propagation in the range from x to $(x+\triangle x)$ meters from the seed source to the length of this range is the product of the total number of seeds and the probability density of the random event distribution function propagating to the distance x meters.

According to Equation. (7), we define $\frac{\varphi(x)}{dx}$ as the seed wind propagation intensity from x to $(x+dx)$ meters from the seed source. Since $\varphi(x)$ is the number of seeds that have settled and accumulated in the range from x to $(x+dx)$ meters from the seed source, we can denote $\frac{\varphi(x)}{dx}$ as $\frac{dQ(x)}{dx}$ and Equation. (7) becomes:

$$\frac{dQ(x)}{dx} = Q \times g(x)$$

9

Equation. (8) reflects the product of the probability density of the distribution function of the random variable distribution function in which the wind propagation intensity is the total number of seeds and the seed wind propagation distance x in the range from x to $(x + dx)$ meters from the seed source. According to this function, we can obtain the probability density of the total number of dandelion seeds and the distribution function of the seed propagation distance, and then determine the intensity of the number of seeds accumulated by sedimentation at this distance.
The relationship between seed wind propagation distance x and wind speed μ is [2-3]:

$$x = \frac{\mu H}{F}$$

10

In this equation, H is the seed release height and F is the seed sedimentation rate, which are constants for specific plants.

Due to the influence of atmospheric turbulence and underlying surface, it can be considered that the wind speed is a continuous random variable[4], and if the wind speed is set to the random variable ξ, then its probability density is $f(y)$ and the distribution function is $F(y)$, and further, the random variable η of the seed wind propagation distance is a function of the random variable ξ of the wind speed, and Equation. (9) turns to:

$$\eta = \frac{\xi H}{F} [10][11]$$

11

According to the calculation method of the probability density function of the random variable function, we have:

$$G(x) = P(\eta \leq x)$$

12

Substitute Equation. (10) into Equation.(11), which is:

$$G(x) = P(\frac{\xi H}{F} \leq x)$$

13

Let:

$$y = \frac{xF}{H}$$

14

Substituting Equation. (13) into Equation. (12), according to the nature of the distribution function of the random variables, then:

$$G(x) = P(\xi \leq \frac{xF}{H}) = P(\xi \leq y) = \int_{-\infty}^{\frac{xF}{H}} f(y)dy$$

15

Since the wind speed is a non-negative value, then Equation. (14) becomes:

$$G(x) = \int_{0}^{\frac{xF}{H}} f(y)dy$$

16

According to the definition of the probability density function of a random variable, then:

$$g(x) = f(\frac{xF}{H}) \times \frac{F}{H}$$

17

Substituting Equation. (16) into Equation. (8), then:

$$\frac{dQ(x)}{dx} = Q \times f(\frac{xF}{H}) \times \frac{F}{H}$$

18

Equation. (17) shows that the intensity of seed wind propagation from x to $(x+dx)$ m from the dandelion seed dispersal source is the product of the total number of seeds as a function of the probability density of wind speed, and the ratio of seed sedimentation rate to sedimentation height.

According to this function, we can know the total number of plant seeds, the seed sedimentation rate, the sedimentation height and the probability density function of the wind speed distribution and then can determine the wind propagation intensity of seeds in different distance ranges, and can calculate the number of seeds that have sedimentation accumulation in the range of propagation distance from x_1 to e.

19

Set ψ to be the natural logarithm of the random variable ξ at wind speed, which is $\psi = \ln\xi$, with a probability density of $h(x)$ and a distribution function of $H(x)$, then:

$$\xi = e^{\psi} \text{ [12]}$$

20

e can be used as 2.71828. Depending on the definition and nature of the probability distribution, we have:

$$F(x) = P(\xi \le y) = P(e^{\psi} \le y) = P(\psi \le \ln y) = \int_0^{\ln y} h(x)dx$$

21

According to the method of calculating the probability distribution density of the random variable function, then:

$$f(y) = h(\ln y) \times \frac{1}{y}$$

22

Set the random variable ψ follow a normal function, which is the probability distribution of wind speed is log-normal[5-6], then the probability density function of ψ is:

$$h(x) = \frac{1}{\sigma_x\sqrt{2\pi}}.\exp(-[\frac{x-\mu_x}{\sqrt{2}\sigma_x}]^2) \text{ [10~12]}$$

23

In this equation, σ_x is the standard deviation of the random variable ψ and μ_x s the mean of the random variable ψ.

According to the characteristics of the random variable function, the wind speed probability density function is:

$$f(y) = \frac{1}{y\sigma_{\ln y}\sqrt{2\pi}}.\exp(-[\frac{\ln y-\mu_{\ln y}}{\sqrt{2}\sigma_{\ln y}}]^2)$$

24

In this equation:

$$\sigma_{\ln y} = \sqrt{\frac{\sum_{i=1}^{n} (\ln y_i - \mu_{\ln y})^2}{n}} \quad [10\sim12]$$

$$\mu_{\ln y} = \frac{1}{n}\sum_{i=1}^{n} \ln y_i = \ln \prod_{i=1}^{n} y_i^{\frac{1}{n}}$$

25

Let

$$\overline{y}_g = \prod_{i=1}^{n} y_i^{\frac{1}{n}}$$

26

So:

$$\mu_{\ln y} = \ln \overline{y_g}$$

27

Substituting Equation. (26) into Equation. (22), then:

$$f(y) = \frac{1}{y\sigma_{\ln y}\sqrt{2\pi}}.\exp(-[\frac{\ln(y/\overline{y_g})}{\sqrt{2}\sigma_{\ln y}}]^2)$$

28

According to the wind speed probability density function, the kernel model of seed propagation is derived from equation (17) as follows:

$$\frac{dQ(x)}{dx} = Q \times \frac{1}{x\sigma_{\ln y}\sqrt{2\pi}}.\exp(-[\frac{\ln(\frac{xF}{Hy_g})}{\sqrt{2}\sigma_{\ln y}}]^2)$$

29

According to Equation. (13), Equation. (28) can become:

$$\frac{dQ}{dx} = \frac{Q}{x\sigma_{\ln y}\sqrt{2\pi}}.\exp(-[\frac{\ln((\frac{xF}{H})/\overline{y_g})}{\sqrt{2}\sigma_{\ln y}}]^2)$$

30

Equation. (29) reflects that the wind propagation intensity of seeds is a function of the probability density of the number of seeds, the distance of seed propagation, and the wind speed distribution.

According to Equation. (29), when the probability density of wind speed is the largest, the strength of the seed wind-borne nucleus will also be the largest. In order to analyze the maximum value of the probability density of wind speed, the following equation is obtained by performing a natural logarithmic transformation on both sides of equation (22) at the same time:

$$\ln f(y) = \ln(\frac{1}{y\sigma_{\ln y}\sqrt{2\pi}}.\exp(-[\frac{\ln y - \mu_{\ln y}}{\sqrt{2}\sigma_{\ln y}}]^2)) = \ln(\frac{1}{y}) + \ln(\frac{1}{y\sigma_{\ln y}\sqrt{2\pi}}) - [\frac{\ln y - \mu_{\ln y}}{\sqrt{2}\sigma_{\ln y}}]$$

31

According to the method of finding the maximum value of the function, find the first derivative of y about both sides of equation (30) at the same time, and make the left value of the equation 0, so:

$$-\frac{1}{y}-2(\frac{\ln y-\mu_{\ln y}}{\sqrt{2}\sigma_{\ln y}})(\frac{1}{y\sqrt{2}\sigma_{\ln y}})=0$$

32

Solving Equation. (31) yields:

$$y=\frac{\exp(\mu_{\ln y})}{\exp(\sigma_{\ln y}^2)}$$

33

Substituting the value calculated in equation (32) into equation (10) yields the distance of seed wind propagation at the maximum probability density, that is, the propagation distance at which the seed propagation density is the largest, and the seed propagation intensity is the greatest at this distance.

The mathematical expectation for the random variable of wind speed in equation (22) is:

$$E(y)=\exp(\mu_{\ln y}+0.5\sigma_{\ln y}^2)$$

34

The variance of the random variable of wind speed in equation (22) is:

$$v(y)=\exp(2\times\mu_{\ln y}+\sigma_{\ln y}^2)\times(\exp(\sigma_{\ln y}^2)\text{-}1)$$

35

Substituting equation (33) into equation (9), the mean distance of seed wind propagation is:

$$\lambda_m=\frac{E(y)H}{F}$$

36

Substituting equation (35) into equation (29), seed propagation intensity at the average propagation distance is yielded.

In this model, the wind speed distribution parameters $\sigma_{\ln y}$ and $\sigma_{\ln y}$ are calculated based on the observed wind speed data using equations (23)-(26), and the Q, H and F in the model method are determined by field investigation and literature review.

Dandelion is a typical plant whose seeds are dispersed by wind[7]. In this paper, six species of dandelion were selected, and their geographical distribution, seed release height H and seed sedimentation rate F were determined through literature review[7][8].

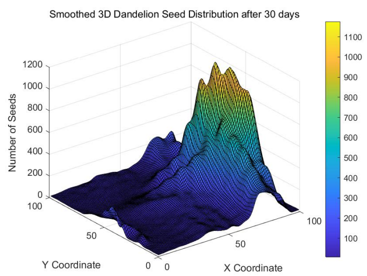
Considering the geographical distribution area of these dandelions [8], we can select three sites where these dandelions are distributed for analysis to compare the differences in the dispersal of dandelion seeds by wind in different locations. In addition, considering that the phenological time of seed dispersal of these plants is mainly from May to October [9], we selected the daily values of wind speed from January to mid-December at three locations as the basic wind element data, and the seed dispersal distance was taken as the spacing in the simulation,with reference[10-12] and the parameters $\sigma_{\ln y}$ and $\mu_{\ln y}$ in the model were calculated according to Equations (23)-(26) (see Table 2).

Month	Mexico Cancún		Manaus Brazil		US. Tucson	
	Longtitude	Latitude	Longtitude	Latitude	Longtitue	Latitude
	86° 50' 47.616" E	21° 10' 27.444" N	*60°01′29″ W*	*3°06′06″ S*	*110°55′35.32″ W*	*32°13′18.27″ N*
	$\overline{y}_g$	σ_{lny}	$\overline{y}_g$	σ_{lny}	$\overline{y}_g$	σ_{lny}
1	8.7	0	2.6	0	8.3	0
2	8.8	0	2.7	0	8.1	0
3	9.2	0	2.6	0	8.3	0
4	8.8	0	2.6	0	8.6	0
5	8.2	0	2.6	0	8.4	0
6	7.7	0	2.8	0	7.8	0
7	6.9	0	2.9	0	6.1	0
8	6.7	0	2.8	0	5.5	0
9	7	0	2.7	0	6.6	0
10	7.9	0	2.5	0	7.5	0
11	8.6	0	2.4	0	8.2	0
12	8.6	0	2.5	0	8.3	0

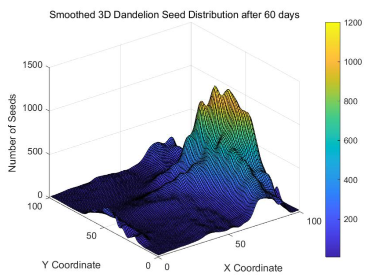
Table 1 Calculation table of $\overline{y}_g$ and σ_{lny} under the obedience of wind speed to lognormal distribution

4.5 Results

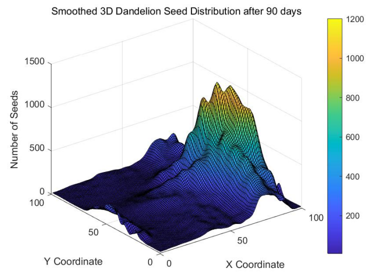
1 Month: The impact of wind direction and wind speed on seed spread is significant.



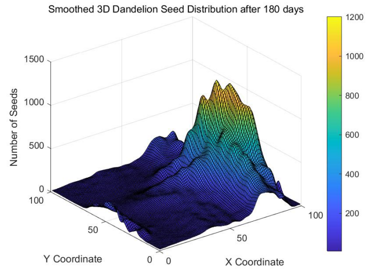
2 Months: The color comparison table on the right clearly shows the growth of the number of dandelions per unit area.



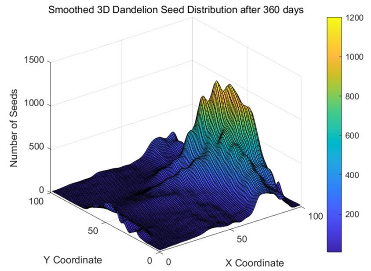
3 Months: The distribution after 3 months is the same as after 6 months, given the flowering time of dandelion



6 Months



12 Months

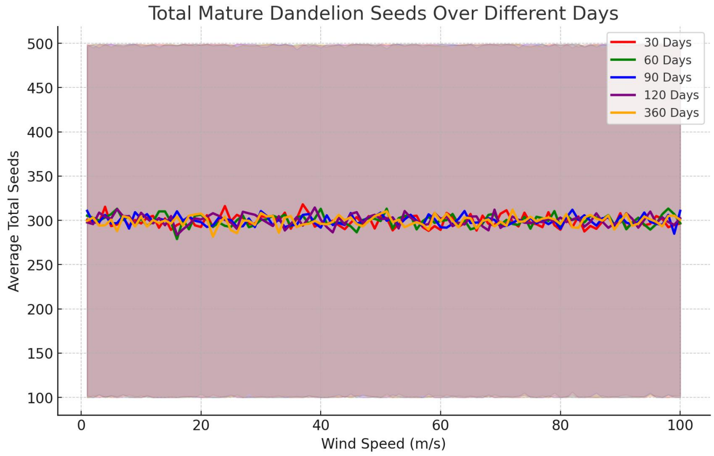


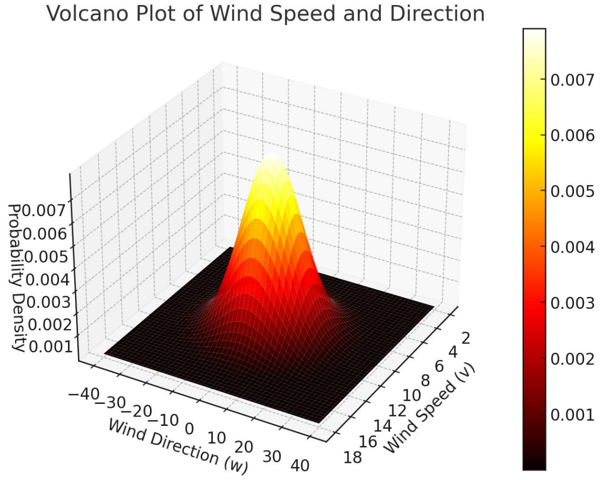
4.6 Sensitivity Analysis

In order to identify which factors always have a significant impact on seed dispersion, and because we consider that factors other than wind speed and direction, such as temperature, humidity and light, merely affect maturity rate, we will only discuss the sensitivity test of wind speed and direction to the total number of seed propagated.

Wind speeds range from 1 m/s to 100 m/s, is corresponding to wind directions from 0 to 360 degrees. The relationship to the total number of seed is shown in the figure below.

No matter how wind speed changes, the average value of wind direction does not change. The seed number always converges to 300 in total and do not exceed the predetermined range, the red interval below which is the total number of dandelions that the land can carry.





4.7 Conclusion

Dandelion Dispersion Model (DDM) with a group of equations, as described in Section A, are created by ourselves innovatively to simulate the change in dandelion population over time.

Typical results showed in Table B and Figure A reflect a pattern of annual periodic change, which is rational and realistic due to the presence of seasonality factors A and B.

All of the arbitrary parameters are analyzed for its sensitivity. The results prove that our model is relatively stable and dependable. All initial values are analyzed repeatedly and thus we are able to know that wind direction and wind speed are the factors which bring the most radical affection on seed dispersion.

Eventually, we reached the conclusion that.

5 Species Invasive Coefficient Evaluation Model(SICEM)

5.1 Variables and Definitions

Variables	Description
RTRM	Reproduction Type (RT) with a focus on its Reproduction Mode (RM) aspect, indicating how the method of reproduction impacts certain outcomes or
LHRM	Life History (LH) evaluated in the context of Reproduction Mode (RM), focusing on how the organism's life cycle stages correlate with its mode of reproduction
PFMRC	Pollination and Fertilization Method (PFM) considered in relation to Reproductive Capacity (RC), examining how these methods impact the
SQRC	Seed Quantity (SQ) analyzed in terms of Reproductive Capacity (RC), focusing on how the number of seeds produced relates to overall reproductive success
SDCRC	Seed Dormancy Characteristics (SDC) in the context of Reproductive Capacity (RC), exploring how seed dormancy affects reproductive potential and success
SSRRC	Seedling Survival Rate Size (SSR) considered with Reproductive Capacity (RC), focusing on how seedling survival impacts reproductive success
SGSRG	Seed Germination Speed (SGS) in the context of Rapid Growth (RG), assessing how quickly seeds germinate and its relation to overall growth speed
PGSRG	Plant Growth Speed (PGS) evaluated with Rapid Growth (RG) in mind, focusing on the rate of plant growth and its implications
GSGC	Genetic Stability (GS) considered in the context of Genetic Characteristics (GC), focusing on how stable genetics affect overall genetic traits
PGC	Polyploidy (P) analyzed in relation to Genetic Characteristics (GC), focusing on the impact of polyploidy on genetic traits
HRSGC	Hybridization with Related Species (HRS) in the context of Genetic Characteristics (GC), examining how hybridization affects genetic traits
$TICA$DEL	Type and Importance of Crops Affected (TICA) evaluated with Direct Economic Loss (DEL) in mind, focusing on how crop types and their importance relate to
AACDEL	Area Affected by Crops (AAC) considered in the context of Direct Economic Loss (DEL), focusing on the relationship between the extent of crop-affected
IYQDEL	Impact on Yield and Quality (IYQ) in relation to Direct Economic Loss (DEL), focusing on how yield and quality changes affect economic losses
IPCIEH	Increase in Production Costs (IPC) evaluated in the context of Indirect Economic Harm (IEH), focusing on how increased production costs contribute to indirect
$IDIM$IEH	Impact on Domestic and International Markets (IDIM) considered in relation to Indirect Economic Harm (IEH), focusing on how market impacts relate to
CQPDDE	Carrying Quarantine Pests (CQP) in the context of Direct Damage to the Ecological Environment (DDE), assessing how quarantine pests contribute to
IBDDE	Impact on Biodiversity (IB) evaluated with Direct Damage to the Ecological Environment (DDE) in mind, focusing on how biodiversity impacts correlate
IEBDDE	Impact on Ecological Balance (IEB) considered in the context of Direct Damage to the Ecological Environment (DDE), focusing on how ecological balance
SEIHE	Soil Erosion (SE) in relation to Indirect Harm to the Ecological Environment (IHE), focusing on how soil erosion contributes to indirect ecological harm
SDIHE	Soil Desertification (SD) evaluated in the context of Indirect Harm to the Ecological Environment (IHE), focusing on how desertification contributes to indirect ecological harm
OAEIHE	Other Adverse Effects (OAE) considered in relation to Indirect Harm to the Ecological Environment (IHE), assessing how various other effects contribute to
PTHH	Plant Toxicity (PT) in the context of Harmfulness to Humans (HH), focusing on how plant toxicity affects human health and safety
SHSHH	Secretion of Harmful Substances (SHS) evaluated in relation to Harmfulness to Humans (HH), assessing how the secretion of harmful substances by plants
DHSH	Degree of Harm (DH) considered in the context of Societal Harmfulness (SH), focusing on the overall harm caused to society
DRDD	Distribution Range (DR) in the context of Domestic Distribution (DD), focusing on how the range of distribution relates to domestic spread
StDD	Status (St) evaluated in relation to Domestic Distribution (DD), focusing on the status of a species or issue in relation to its domestic distribution
DRID	Status (St) evaluated in relation to International Distribution (ID), focusing on the status of a species or issue in relation to its domestic distribution
StID	Distribution Range (DR) in the context of International Distribution (ID), focusing on the international aspect of distribution
HTDHC	Habitat Type Diversity (HTD) evaluated in relation to Habitat Characteristics (HC), focusing on how diverse habitat types relate to overall habitat
WMSNS	Wind-Mediated Spread (WMS) in the context of Natural Spread (NS), assessing how wind contributes to natural dispersion
WFSNS	Water Flow Spread (WFS) considered in relation to Natural Spread (NS), focusing on water flow's role in natural dispersion
AMSNS	Animal-Mediated Spread (AMS) in the context of Natural Spread (NS), focusing on how animals contribute to natural dispersion
BENS	Border Entry (BE) evaluated with Natural Spread (NS) in mind, assessing how crossing borders contributes to natural spread
IIHMS	Intentional Introduction (II) in the context of Human-Mediated Spread (HMS), focusing on how deliberate introductions contribute to human-mediated
UIHMS	Unintentional Introduction (UI) considered in relation to Human-Mediated Spread (HMS), assessing how accidental introductions contribute to human-
RLKIID	Required Level of Professional Knowledge (RLK) in the context of Difficulty of Inspection and Identification (IID), focusing on how the required professional
PSPIID	Presence of Similar Plants in Our Country (PSP) evaluated in relation to Difficulty of Inspection and Identification (IID), focusing on how the presence of
PCEDC	Physical Control Effectiveness (PCE) in the context of Difficulty of Control (DC), assessing how effective physical control methods are in managing the issue
CCEDC	Chemical Control Effectiveness (CCE) considered in relation to Difficulty of Control (DC), focusing on the efficacy of chemical control methods
BCEDC	Biological Control Effectiveness (BCE) in the context of Difficulty of Control (DC), assessing the effectiveness of biological control strategies
DLED	Difficulty Level (DL) in the context of Eradication Difficulty (ED), focusing on the overall challenge level in eradicating the issue
DCSDCS	Domestic Climate Suitability (DCS) evaluated in the context of itself (DCS), focusing on how suitable the domestic climate is for the issue at hand
ECEC	Environmental Conditions (EC) considered in relation to itself (EC), assessing the overall environmental conditions affecting the issue
STST	Stress Tolerance (ST) in the context of itself (ST), focusing on how tolerance to various stresses impacts the issue
$PDNE$DNE	Presence of Domestic Natural Enemies (PDNE) in the context of itself (DNE), assessing how the presence of natural enemies within the domestic environment

5.2 TOPSIS

The TOPSIS method determines the optimal solution based on the distance of each alternative relative to the ideal and negative ideal solutions.

1

Let R(m,n) be the original decision matrix, m be the number of evaluation objects, and n be the number of indicators. The normalized decision matrix R'(m,n) is:

$$R'_{ij} = \frac{R_{ij}}{\sqrt{\sum_{i=1}^{n} R_{ij}^2}} \quad for \quad i = 1,2,\dots,n; \quad j = 1,2,\dots,m$$

2

Calculate the weighted normalized decision matrix. Let W_n be the weight of each evaluation index, then the weighted normalized decision matrix V(m*n) is:

$$V_{ij} = W_n \times R'_{ij} \quad for \quad i = 1,2,\dots,n; \quad j = 1,2,\dots,m$$

3

Determine the positive and negative ideal solutions. The formulas for positive ideal solution A+ and negative ideal solution A- are :

$$A^+ = \{\max(V_{ij}) \mid j \in J\}, \{\min(V_{ij}) \mid j \in J'\} \quad for \quad i = 1,2,\dots,m$$

$$A^- = \{\min(V_{ij}) \mid j \in J\}, \{\max(V_{ij}) \mid j \in J'\} \quad for \quad i = 1,2,\dots,m$$

J is the set of equity metrics and J' is the set of cost metrics here.

4

Calculate the distance between each evaluation object and the positive and negative ideal solutions. The distance S+_i from the positive ideal solution and the distance S-_i from the negative ideal solution are:

$$S_i^+ = \sqrt{\sum (A_j^+ - V_{ij})^2} \quad for \quad i = 1,2,\dots,m$$

$$S_i^- = \sqrt{\sum (A_j^- - V_{ij})^2} \quad for \quad i = 1,2,\dots,m$$

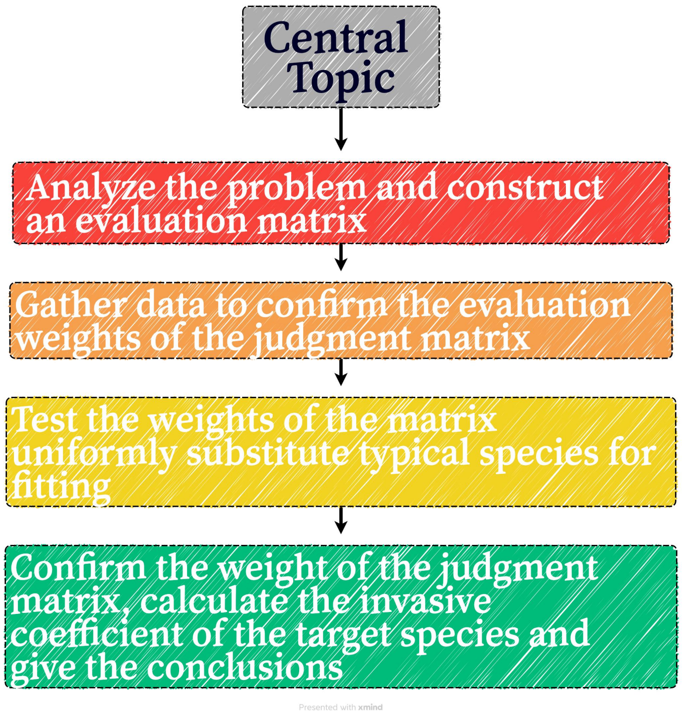
5

Calculate the composite score of each evaluation object. The equation for calculating the composite score C*_i of each evaluation object is:

$$C_i^* = \frac{S_i^-}{S_i^+ + S_i^-} \quad for \quad i = 1,2,\dots,m$$

The object with the highest score is the optimal solution.

5.3. AHP



Analytic Hierarchy Process, known as AHP, is a model that aids people in making decisions by applying pairwise comparisons. Linear algebra is the basic principle of this mathematical thinking process. Fundamentally, priorities for alternatives and the criteria used for judging the alternatives are developed in the AHP. When we are intended to use the AHP, it is necessary for us to construct a hierarchic or a network structure to represent that problem and use pairwise comparisons to establish relations within the structure. With AHP, the indexes we take into consideration will be ranged in hierarchic structure.

The process is divided into two steps. Our first step is to construct a hierarchical model.

The mutual relations, the decision objectives, factors and decision objects are divided into the highest level, the very strong high level, the slightly strong level, the middle level, the slightly slow level, the very slow level and the lowest level separately.

Numerical scale	The level of importance
1	The most significant level
2	The extremely significant level
3	The slightly significant level
4	Moderate Level
5	The slightly insignificant level
6	The extremely insignificant level
7	The most insignificant level

Next we are going to construct a judgment matrix.

$$[a_{ij}] = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix}$$

Data analysis procedure are following steps.

First, the pairwise comparison matrix, matrix A, is extracted from data of previous essays.If the formula 7.8 is not confirmed for all k, j, and i, the eigenvector method will be applied.

$$a_{ik} \cdot a_{kj} = a_{ij}$$

If the matrix is incompatible and incomplete consistency exists, the pair comparisons latrix cannot be used normalizing column to get W_i. For a positive and reversed matrix, the eigenvector technique can be applied based on the llowing two formulas:

$$e^T = (1,1,\dots,1)$$

$$W = \lim_{k\to\infty} \frac{A^k \cdot e}{e^T \cdot A^k \cdot e}$$

When matrix is incompatible, calculation should be repeated several times to make a decision to reach a convergence among the set of answers which is involved in a successive repetition of this process.

Second step is using the score of importance table provided by [文章] We have to evaluate and judge the consistency of the judgment matrix. It'shas the following properties:

$$a_{ij} = \frac{1}{a_{ji}}$$

$$Aw = \lambda_{max} w, \lambda_{max} \geqslant n$$

$$\lambda_{max} = \frac{\sum a_j w_j - n}{w_1}$$

Then, these three formulas given above are used to transform the raw data into meaningful values.

The consistency index is defined as:

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

Our next step is to calculate the consistency ratio (CR) in order to confirm the results from AHP. If CR < 0.10, the agreement test is passed, otherwise the judgment matrix needs to be reconstructed.

The formula is shown below:

$$CR = \frac{CI}{RI}$$

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

$$RI = \frac{CI_1 + C_2 + \dots + CI_n}{n}$$

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

	Invasiveness	Adaptability	Spreadability	Harmfulness	
RI	1.61	0.9	1.51	1.56	
CR	1.38E-16	0	0	0	3.45E-17

Since CR<1, Consistency test is valid.

5.4 Results&Conclusion

In this section, we will elaborate how our model, in which invasiveness is most significant, apply for a typical species, and we will show how our model fits the preference of region and the special needs of species.

1.All the scales considering our indexes can be divided into 7 degrees(1,1,5,2,2.5,3,4,,5scores)

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

$$a_{ij} = \frac{1}{n}, n \in N^*$$

Thus, the fifinal judgment matrix of this student will be few parts:

$$\text{juece matrix}=\begin{bmatrix} 1.0000 & 4.0000 & 1.6000 & 1.6000 \\ 0.2500 & 1.0000 & 0.4000 & 0.4000 \\ 0.6250 & 2.5000 & 1.0000 & 1.0000 \\ 0.6250 & 2.5000 & 1.0000 & 1.0000 \end{bmatrix}$$

First Sub-Matrix (Columns 1 to 8)

$$\begin{bmatrix} 1.0000 & 1.0000 & 2.0000 & 0.6667 & 0.4444 & 1.0000 & 0.5714 & 1.0000 \\ 1.0000 & 1.0000 & 2.0000 & 0.6667 & 0.4444 & 1.0000 & 0.5714 & 1.0000 \\ 0.5000 & 0.5000 & 1.0000 & 0.3333 & 0.2222 & 0.5000 & 0.2857 & 0.5000 \\ 1.5000 & 1.5000 & 3.0000 & 1.0000 & 0.6667 & 1.5000 & 0.8571 & 1.5000 \\ 2.2500 & 2.2500 & 4.5000 & 1.5000 & 1.0000 & 2.2500 & 1.2857 & 2.2500 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Second

$$\begin{bmatrix} 1.6000 & 1.6000 & 1.6000 & 1.6000 & 1.0000 & 0.3636 & 0.8000 & 0.8000 \\ 1.6000 & 1.6000 & 1.6000 & 1.6000 & 1.0000 & 0.3636 & 0.8000 & 0.8000 \\ 0.8000 & 0.8000 & 0.8000 & 0.8000 & 0.5000 & 0.1818 & 0.4000 & 0.4000 \\ 2.4000 & 2.4000 & 2.4000 & 2.4000 & 1.5000 & 0.5455 & 1.2000 & 1.2000 \\ 3.6000 & 3.6000 & 3.6000 & 3.6000 & 2.2500 & 0.8182 & 1.8000 & 1.8000 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Third

$$\begin{bmatrix} 0.5000 & 1.0000 & 0.8000 & 1.3333 & 1.3333 & 2.0000 \\ 0.5000 & 1.0000 & 0.8000 & 1.3333 & 1.3333 & 2.0000 \\ 0.2500 & 0.5000 & 0.4000 & 0.6667 & 0.6667 & 1.0000 \\ 0.7500 & 1.5000 & 1.2000 & 2.0000 & 2.0000 & 3.0000 \\ 1.1250 & 2.2500 & 1.8000 & 3.0000 & 3.0000 & 4.5000 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Given the large size of the matrix `data_matrix`, we will represent it as a series of sub-matrices, each corresponding to a portion of the columns. Here's how it can be presented using mathematical notation:
First Sub-Matrix (Columns 1 to 8):

$$\begin{bmatrix} 1.0000 & 1.0000 & 1.0000 & 1.3333 & 1.3333 & 1.3333 & 1.3333 & 0.4000 \\ 1.0000 & 1.0000 & 1.0000 & 1.3333 & 1.3333 & 1.3333 & 1.3333 & 0.4000 \\ 1.0000 & 1.0000 & 1.0000 & 1.3333 & 1.3333 & 1.3333 & 1.3333 & 0.4000 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & & & & \end{bmatrix}$$

2. **Second Sub-Matrix (Columns 9 to 16):**

$$\begin{bmatrix} 0.5000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 & 1.0000 & 0.5000 & 0.8000 \\ 0.5000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 & 1.0000 & 0.5000 & 0.8000 \\ 0.5000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 & 1.0000 & 0.5000 & 0.8000 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & & & & \end{bmatrix}$$

3. **Third Sub-Matrix (Columns 17 to 24):**

$$\begin{bmatrix} 0.8000 & 0.8000 & 0.8000 & 1.0000 & 1.0000 & 0.6667 & 0.6667 & 0.8000 \\ 0.8000 & 0.8000 & 0.8000 & 1.0000 & 1.0000 & 0.6667 & 0.6667 & 0.8000 \\ 0.8000 & 0.8000 & 0.8000 & 1.0000 & 1.0000 & 0.6667 & 0.6667 & 0.8000 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & & & & \end{bmatrix}$$

These matrices are parts of the larger judgment matrix. To visualize the entire matrix, you would concatenate these sub-matrices horizontally in the order they are presented. The `\vdots` indicates that the pattern continues down the matrix for the rows that were not fully displayed.

Finally, we put the weighting into the TOPSIS system to rank the score of invasive species.

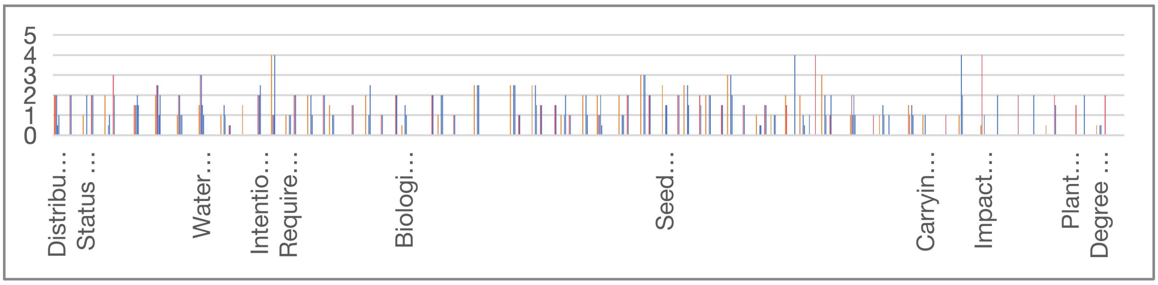
According to the judgement matrix above, we are able to substitute the data of Dandelion.

$$C_i^* = \frac{S_i^-}{S_i^+ + S_i^-} \quad for \quad i = 1, 2, \ldots, m$$

The score of Dandelion is 69.

5.5 Sensitivity Analysis

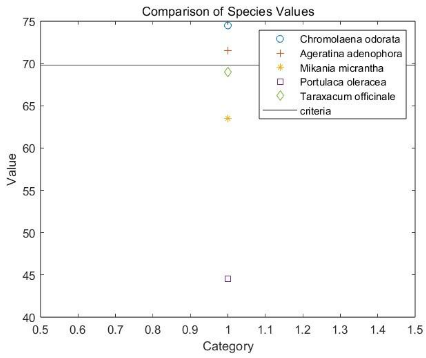
To test the accuracy of our model and to discuss the evaluation criteria for invasive species, we bring five typical species back into the model, calculate their scores, and match them with realistic evaluation results



C.odorata	A.adenophora	M.Kunth	P.oleracea	T.officinale	Average Value of Invasive Species
74.5	71.5	63.5	44.5	69	69.83333333

5.6 Result

According to the rating and evaluation criteria of dandelion above, we use the error bar, and the invasive species standard line is within the range of the dandelion error bar, so the dandelion is an invasive species



6 Application

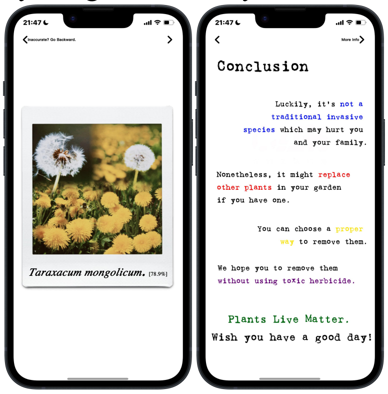
We develop an App for individuals to identify invasive species. After entering into the App, user is required to choose the region they are in. Our app can locate them automatically if they allow.

Users are required to set their location first or allow us to locate them automatically for more precise location. Then, the climate data, especially the wind direction and wind speed are derived through using weatherspark.com according to the information provided previously.

Next, customers can take pictures of a plant and upload. We suggest them to provide us with more than one picture with high clarity and close-up details; thus, the result derived from Pl@nt net, one of the most authoritative plant identifiers, will be more accurate since plants in same genus are extremely similar in many aspects. Consequently, we will also give probability of being the right species and a tip “Just for your information”

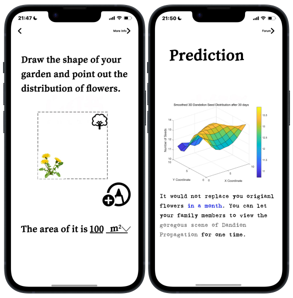
There will be 3 kinds of results. 「Result 1: The plant species is definitely an Invasive Species according to present criteria of identifying Invasive species.」

「Result 2: The plant species is not an Invasive species according to presnent criteria of identifying invasive species. However, it spreads fast, grow strong and might replace the flowers in your garden or if you have bees.」



「Result 3: The plant species is not an Invasive species according to present criteria of identifying invasive species. It won’t have any effect on you. We suggest you to save them since “Plants live matter.”」

What’s more, if is 「Result 2」. The user can choose to add more information, like the shape of his/her garden and the distribution of present flowers. According to the Dandelion Dispersion Model (DDM), a picture of distribution of all species will be provided combined with the land shape. We choose to do this because we consider a circumstance which might happen in real life: a family has a garden. The mom likes to plant various beautiful flowers in the garden. However, one day, dandelion comes. If the mom uses our App, she will know that those dandelions might replace all the flowers she makes an effort to foster; then, she might decide to clean those dandelions. Nonetheless, if her kids would like to see how dandelion weeds disperse for one time, the mom will be in conflict. Fortunately, as long as she provides us with the present distribution of plants in her garden, we will come up with a predicting distribution in 1 month and 3 months to tell the mom whether dandelion will affect her flowers when dandelion first reproduce. If dandelion reproduction won't affect those flowers after their first reproduction, the mom can choose to leave those dandelion for a month, waiting for them first reproduction and then remove them; therefore, both of their desires can be achieved.



The purpose of this application aims to help individuals to understand specific plant species better and create a peaceful environment between people and plants. Mathematical model could be merely quantitative, but it could be utilized with human touch.

7 Strengths and Limitations

7.1 Limitations

Question 1

- We do not consider the effects native species, especially natural predators of dandelion, could bring; thus, results are not that authentic according to the reality though dandelion can replace most of other plant species.
- Certain factors, such as soil fertility and water occupancy, are hard to convert into quantitative data; consequently, results might be slightly deviate from the real situation.
- We eliminate the effect which emergent cases will bring, such as much farther distance and larger wind speed caused by storms; nonetheless, those data could be considered as outliers.

Question2

- Using AHP might cause subjective component accounts for a large proportion when evaluating a specific species; thus, it might affect the result more or less.
- AI identification might get wrong when there are 2 species extremely similar.

7.2 Strength

Question1

- We take the variation of climate in different regions into account, especially the wind speed; therefore, the result will be more realistic.

- We build an innovative model which is applicable and sufficient to simulate the dandelion dispersion in a year for most situations.
- Results derived match our model perfectly. Apart from that, our model also performs well in comparison with other models,

Question2

- Using TOPSIS enhances the objectivity of our model effectively.
- The design and combination of models and application make those models more beneficial and practical for people, rather than merely finishing a problem.
- Results derived correspond to our model well and the data fitting is very successful, which proves the accuracy of the model.

References

[1] Andersen M C.Mechanistic models for the seed shadows of wind dispersed plants[J].AmericanNaturalist,1991,137(4):476-497.
[2] Nathan R,Katul G G,Horn H S,et al.Mechanisms of long-distance dispersal of seeds by wind[J].
[3] Nathan R,Schurr F M,Spiegel O,et al.Mechanisms of long-distance seed dispersal[J].Trends in Ecology and Evolution,2008,23(11):638-647.
[4] Che Junhui, Zhao Ping, Shi Qian, et al. Advances in Atmospheric Boundary Layer Research [J]. Chinese Journal of Geophysics
[5] Li Hui, Sun Hongbin, Zhang Fang. A Review of Wind Speed Distribution Models in Wind Farms [J]. Advanced Technology of Electrical Engineering and Energ
[6] Wang Miao, Zeng Lihua. Research on Wind Speed Frequency Distribution Models [J]. Journal of Hydroelectric Engineering
[7] Zhang Jian, Zhou Cunyu, Fei Yongjun. Study on the Dispersal Ability of Seeds from Six Dandelion Species [J]. Seed
[8] Ge Xuejun, Lin Yourun, Zhai Datong. Preliminary Arrangement of Dandelion Genus Plants in China [J]. Bulletin of Botanical Research
[9] Wang Dongni, Ren Hongling, Guo Qiao, et al. Phenological Characteristics of Dandelions in Jilin Province and Their Response to Global Warming [J]. Journal of Meteorology and Environment
[10] Li Hui, Sun Hongbin, Zhang Fang. A Review of Research on Wind Speed Distribution Models in Wind Farms [J]. Advanced Technology of Electrical Engineering and Energy
[11] Wang Miao, Zeng Lihua. Research on Wind Speed Frequency Distribution Models [J]. Journal of Hydroelectric Engineering
[12] Yu Yang, Sun Yuejing. Maximum Likelihood Estimation of Parameters for Log-Normal Distribution [J]. Journal of Jiujiang University: Social Science Edition
[13] Roberts, H. F. (1936). Seed reproduction in the dandelion. Scientific Agriculture, 17(4), 235-242.
[14] Fang Yaqin. (2022). Prediction and Analysis of the Spreading Trends of Three Invasive Compositae Plants in China Based on the Ensemble Model (EM) (Master's thesis, Shaanxi Normal University)
[15] Ding Hui, Shi Biqing, & Xu Haigen. (2011). Risk Assessment Index System and Evaluation Methods for Alien Species. Journal of Ecology and Rural Environment
[16] Li Xiangqin, Wei Chunqiang, Wang Yunbo, Tang Saichun, & Pan Yumei. (2017). Risk Assessment of the Invasive Plant Mikania micrantha. Forestry Science and Technology
[17] Xia Guangxu, Cheng Kai. (2017). Risk Analysis of the Harm Caused by the Introduction of Eupatorium adenophorum into Hezhang County. Journal of Anhui Agricultural Sciences