Dandelions:
Friend? Foe? Both? Neither?

Summary Sheet

The problem requests a prediction of the spread from a single dandelion over a 12-month period, under varying weather conditions. It is also required to design an impact factor for identifying invasive species.

To begin with, we divide the life cycle of a dandelion into four stages. We incorporate the ecological statistics into a compact Leslie matrix to model the transition between stages. A discrete-time method is then developed to estimate the population growth in each stage. To visualize the spatial distribution of the dispersed seeds, gamma distribution and uniform distribution are used to model the dispersal distance and direction of each seed. Monte Carlo method is utilized to acquire an accurate distribution of seed location, which is shown to be rotationally symmetric and thus suggests an identical pattern in any direction. The spread distance is likely to follow a gamma distribution, which is validated by a χ^2 goodness-of-fit test.

Two tropical regions are selected to study the seed propagation in more realistic weather conditions. Mumbai is a typical example with distinct dry/wet seasons and consistent wind directions. Rio de Janeiro exemplifies regions with abundant rainfall and evenly distributed wind directions. Based on historical climatic data, the shape parameter of gamma distribution is calculated for dry and rainy season, respectively. Normal distribution is used to model the wind with a dominant direction. The simulation shows that rainy condition hinders the seed from travelling longer distance, while consistent wind direction facilitates the spread as both the mean and maximum dispersal distance are increased.

We also develop an assessment tool for quantifying the negative impact of an invasive plant on the environment. A total of seven criteria are proposed to evaluate different aspects including the impact on the ecosystem, the species' ecological features, and difficulty in management. For each criterion, we devise a set of levels together with detailed descriptors. The overall impact factor is a weighted sum of the points achieved in all the criteria. Tests on both invasive and non-invasive plants show that the impact factor exhibits distinctive values for the two categories, which enables an easy identification of invasive species.

Finally, it is worth noting that the methods proposed in this work are scalable and adaptable, in the sense that they can be applied to other scenarios by adjusting the model parameters. For instance, biological features of a different species can be accommodated by changing the parameters related to the seed number, survival rate, growth period, etc. Probability distribution of the seed dispersal distance and direction can be customized when the local climatic data are available. Given more expertise in biology, the impact factor can also be improved by refining the relative weights and developing more accurate specification for each level.

Table of Contents

1 Introduction

Dandelions are a large genus of flowering plants that originate from Europe. After being introduced into North America, they have propagated widely as wildflowers across temperate regions of the continent. Dandelions thrive in diverse habitats from yards and gardens to roadsides, croplands, and more.

After flowering, a dandelion drops off dried petals and stamens, and the parachute ball opens up into a full sphere. When development is complete, the mature seeds are attached to white, fluffy "parachutes". This is called the "puffball" state. The "parachutes" easily detach from the seed head and glide by wind. The seeds are able to cover large distances when dispersed. Dandelions offer nutritional value. The entire plant is edible, providing high amounts of vitamins A, C, K and moderate sources of calcium, potassium, iron, and manganese. They also provide food for animals such as rabbits, wild turkeys, and birds [1].

However, the same traits of being resilient also make dandelions invasive. In general, invasive species have the ability to adapt, create new populations, and impact the environment, health, and the economy. They compete for food and space, reduce biodiversity in the ecosystem, or cause water shortages. They could also impact the fishing, livestock breeding, and crop cultivation, which hinders economic growth [2].

In this sense, dandelions are invasive, as they aggressively crowd out more desirable species across North America. In some jurisdictions, the species is listed as a noxious weed. Denali National Park and Preserve in Alaska lists the plant as the most common invasive species in the park and hosts an annual Dandelion Demolition event to remove the plant from roadsides [3]. Dandelions are easily dispersed to gardens by their fluffy seeds. Once established, they are extremely difficult to eliminate due to the persistent root systems. Dandelions are weeds, and take up land in lawns and gardens that could be better served by planting native plants that support native bees [3].

To investigate how dandelions are invasive from a mathematical perspective, we shall start with modeling its spread to see how a single dandelion reproduces over a year. Then, we will develop a quantitative tool to evaluate the extent to which dandelions can be invasive to the environment, and compare with other invasive and non-invasive plants.

2 Modeling the spread of dandelions generated by a single seed head

2.1 Key terms and assumptions

A single dandelion can produce numerous seeds when it grows into a seed head (also called puffball). According to the research [1,4,5], a dandelion in the seed head stage can produce about 54-250 seeds, among which about 2 to 4 percent will sprout, blossom, and finally develop into a seed head, ready to propagate its seeds of next generation. It takes a seedling about 2 to 3 months to blossom, and another 2 weeks to form a seed head [6]. After being detached by a gust of wind, the seed head releases the seeds into the air. According to the study [7], more than 99.5% of the seeds travel within 10 meters of their parent. Only 0.05% of the seeds can be dispersed beyond 100 meters.

In order to mathematically model the spread of dandelions, we define the key terms and impose several assumptions as follows.

- A seed head produces N_{seed} seeds, each of which has a probability P_{seed} to be successfully developed. Thus, the number of developed seedlings in the next generation is

  $$N = N_{seed}\, P_{seed}$$

  For the simulations in the subsequent section, we use $N_{seed} = 150,\ P_{seed} = 0.03$.These values yields $N \approx 5$, meaning a seed head produces 5 seedlings in the next generation.
- Under normal weather conditions, the dispersal distance R is a random variable with a mean of μ_R meters.
- The direction in which the seed is propagated is represented by an angle Θ, which is the angle of counterclockwise rotation from due east. The random variable Θ is assumed to follow a uniform distribution from 0 to 360 degrees.

  $$\Theta \sim U(0, 360)$$
- The entire life cycle L from seedling to a mature seed head takes either 3 or 4 months. Thus, the life cycle L of a newly developed seedling only has two possible values. The probability distribution of L is

l	$P(L = l)$
3	P_{short}
4	$1 - P_{short}$

  P_{short} is the probability that the life cycle of a seed is shorter, i.e., $L = 3$. In the subsequent simulations, we use $P_{short} = 0.25$, assuming one fourth of the seedings has accelerated life cycle.

2.2 Estimating the population growth using Leslie matrix

The dandelion life cycle from seedling to mature seed head spans 3-4 months. We can divide this 12-month period into 4 states: S_0, S_1, S_2, B. S_0 represents a new seedling. S_1 and S_2 represent seedlings grown for 1 and 2 months, respectively. B denotes a blossomed flower that will transition into the puffball stage to propagate seeds in the next month, completing the cycle. The transition between these four states is illustrated in Figure 1.

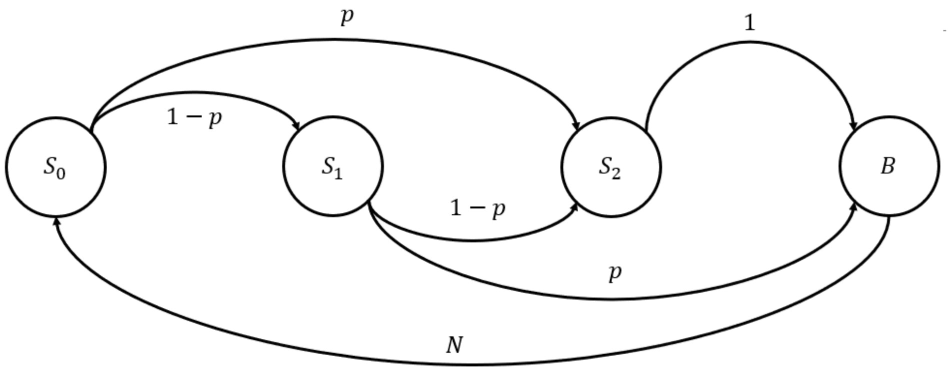


Figure 1 Transition diagram of the four states of a dandelion

The information in the transition diagram can be integrated into a transition matrix **T**. This matrix is also known as the Leslie matrix [8], and it is widely used to represent the transition between different age groups of a species.

$$\mathbf{T} = (t_{ij}) = \begin{pmatrix} 0 & 0 & 0 & N \\ 1-p & 0 & 0 & 0 \\ p & 1-p & 0 & 0 \\ 0 & p & 1 & 0 \end{pmatrix}$$

The first row represents the "fecundity" of each state. Since only a blossom can produce a new seedling, the only non-zero element is $t_{14} = N$, indicating a blossom can produce N alive seedlings in the next month, after taking into account the total number of seeds in a puffball and the survival rate of them.

The first column of **T** shows that a proportion of $(1-p)$ of the seedlings in state S_0 grow into state S_1, while a proportion p of them are advanced into S_2 in the next month. Similarly, the second column indicates the same proportion of dandelions in state S_1, which directly develop into blossom B and skip the state S_2. The only non-zero element in the third row $t_{43} = 1$ means that a dandelion in state S_2 is certain to grow into a blossom. This certainty guarantees that a seedling can accelerate at most once during its life cycle, either from S_0 to S_1 or from S_1 to S_2, so that the minimum life cycle is 3 months while the normal life cycle is 4 months. The overall probability that a seedling has life cycle of 4 months is $(1-p)^2$, and the probability of an accelerated life cycle is

$$P_{short} = 1 - (1-p)^2$$

Therefore, the value of p in matrix **T** can be calculated as

$$p = 1 - \sqrt{1 - P_{short}}$$

In our simulation, $p = 1 - \sqrt{1 - 0.25} \approx 0.134$.

The state matrix $\mathbf{s}_n$ is a column matrix showing the number of dandelions in each state n months after the spread of the initial seed head.

$$\mathbf{s}_n = \begin{pmatrix} q_{0,n} \\ q_{1,n} \\ q_{2,n} \\ q_{b,n} \end{pmatrix}$$

where $q_{0,n}, q_{1,n}, q_{2,n}, q_{b,n}$ represents the amount of S_0, S_1, S_2, B after n months respectively.

The state matrices of two consecutive months can be related by the matrix multiplication

$$\mathbf{s}_{n+1} = \mathbf{T}\mathbf{s}_n$$
$$= \begin{pmatrix} N \cdot q_{b,n} \\ (1-p) \cdot q_{1,n} \\ p \cdot q_{0,n} + (1-p)q_{1,n} \\ p \cdot q_{1,n} + (1-p)q_{2,n} \end{pmatrix}$$

Assume the parent seed head propagates its seeds at the beginning of a 12-month period. The initial state is

$$\mathbf{s}_0 = \mathbf{T}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} N \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

meaning that a total of N seedlings are present initially.

Then the state matrix after n months can be derived by repeatedly pre-multiplying the Leslie matrix $\mathbf{T}$

$$\mathbf{s}_n = \mathbf{T}^n\mathbf{s}_0$$

Substituting the values $p = 1 - \sqrt{0.75}$, $N = 5$, we obtain

$$\mathbf{s}_1 = \mathbf{T}^1\mathbf{s}_0 = \begin{pmatrix} 0 \\ 4.33 \\ 0.67 \\ 0 \end{pmatrix}, \quad \mathbf{s}_2 = \mathbf{T}^2\mathbf{s}_0 = \begin{pmatrix} 0 \\ 0 \\ 3.75 \\ 1.25 \end{pmatrix}, \quad \mathbf{s}_3 = \mathbf{T}^3\mathbf{s}_0 = \begin{pmatrix} 6.25 \\ 0 \\ 0 \\ 3.75 \end{pmatrix}$$

$$\mathbf{s}_6 = \mathbf{T}^4\mathbf{s}_0 = \begin{pmatrix} 7.81 \\ 0 \\ 14.06 \\ 9.38 \end{pmatrix}, \quad \mathbf{s}_9 = \mathbf{T}^6\mathbf{s}_0 = \begin{pmatrix} 9.77 \\ 60.89 \\ 44.58 \\ 17.58 \end{pmatrix}, \quad \mathbf{s}_{12} = \mathbf{T}^{12}\mathbf{s}_0 = \begin{pmatrix} 275.88 \\ 228.35 \\ 101.24 \\ 29.3 \end{pmatrix}$$

The final total number of dandelions of four states is

$$275.88 + 228.35 + 101.24 + 29.3 \approx 635$$

which indicates a single seed head can produce a total of 635 seeds after 12 months. Our simulation result is comparable with the reported statistic in [1, 4, 5], taking into account a survival rate of approximately 2-4%.

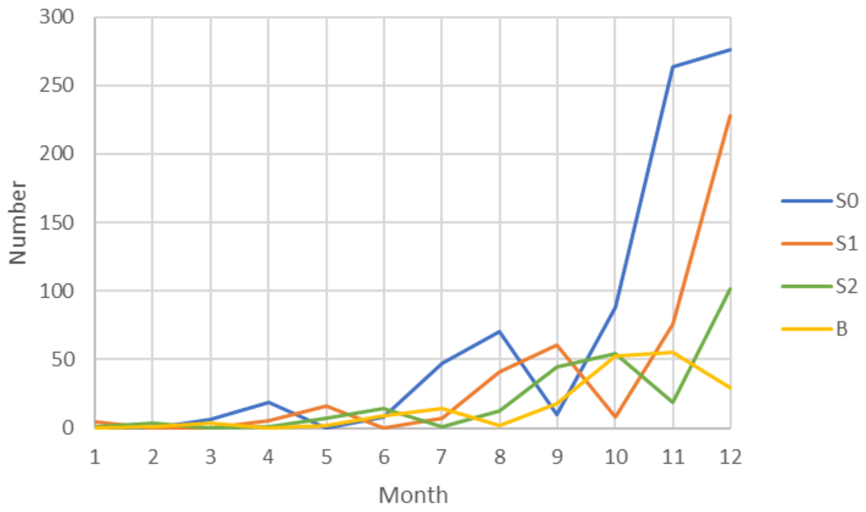


Figure 2 Growth of the four states over 12 months

Figure 2 shows the growth of each state within the 12-month period. There is an evident increase every 3-4 months, which is aligned with the typical life cycle of a dandelion.

2.3 Simulating the spatial propagation of dandelions by a single seed head

Simulation using Leslie matrix only provides the population of the dandelions of each state. To visualize the spatial distribution of the dandelions, we are going to calculate the coordinates of the dandelions and draw a scatter plot of them.

Suppose the parent seed head is at the position (X_0, Y_0), the coordinates of its seed is determined by the dispersal distance R meters and the angle Θ degrees measured from the positive x-axis in the anticlockwise direction, as illustrated in Figure 3.

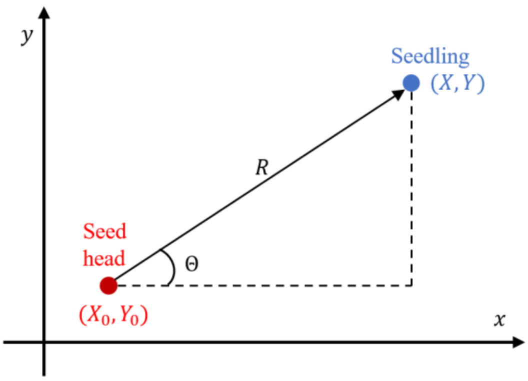


Figure 3 Relation of the coordinates of a parent seed head and its seed

The seed position (X, Y) is calculated using trigonometry

$$\begin{cases} X = X_0 + R\cos\Theta \\ Y = Y_0 + R\sin\Theta \end{cases}$$

2.3.1 Probability distribution of R and Θ

Because the dispersal distance varies from seed to seed, R is considered as a random variable. Normal distribution [9] seems an intuitive model for R.

$$R \sim N(\mu_R,\ \sigma_R^2)$$

where μ_R is the average dispersal distance, and σ_R is the standard deviation. Using the standard normal distribution $Z \sim N(0, 1^2)$, we write the cumulative probabilities reported in [7] as

$$P(R < 10) = P\left(Z < \frac{10 - \mu_R}{\sigma_R}\right) = 0.995$$

$$P(R < 100) = P\left(Z < \frac{100 - \mu_R}{\sigma_R}\right) = 0.9995$$

By applying the inverse of a normal cumulative function, we obtain the system of equations

$$\begin{cases} 10 - \mu_R = 2.575 \cdots \sigma_R \\ 100 - \mu_R = 3.290 \cdots \sigma_R \end{cases}$$

The solution for the mean value μ_R is

$$\mu_R \approx -314$$

which is not sensible, since the average dispersal distance cannot be far below zero.

Another reasonable probability distribution for the dispersal distance is a gamma distribution [9].

$$R \sim \Gamma(\alpha,\ \beta)$$

where α is the shape parameter, β is the scale parameter, and the probability distribution function is

$$f(r\,;\ \alpha,\ \beta) = \frac{1}{\beta^\alpha\ \Gamma(\alpha)}\ x^{\alpha-1}\ e^{-\frac{x}{\beta}}$$

To yield $P(R < 10) = 0.995$, we select the standard gamma distribution, so $\beta = 1$ (the selection of β will be discussed later). Then, a numerical method is adopted to estimate the value of α.

$$\alpha = 3.42$$

Figure 4 shows the probability density function of the resultant gamma distribution.

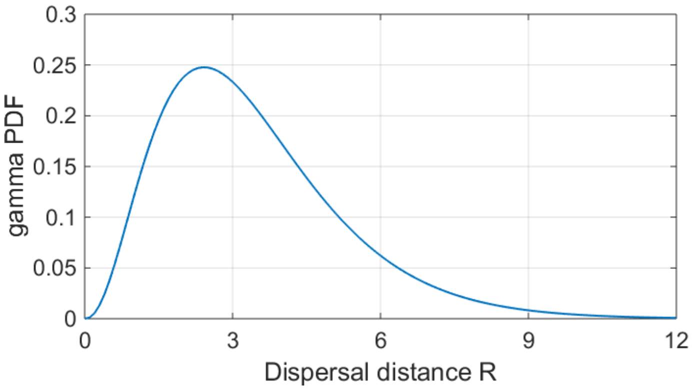


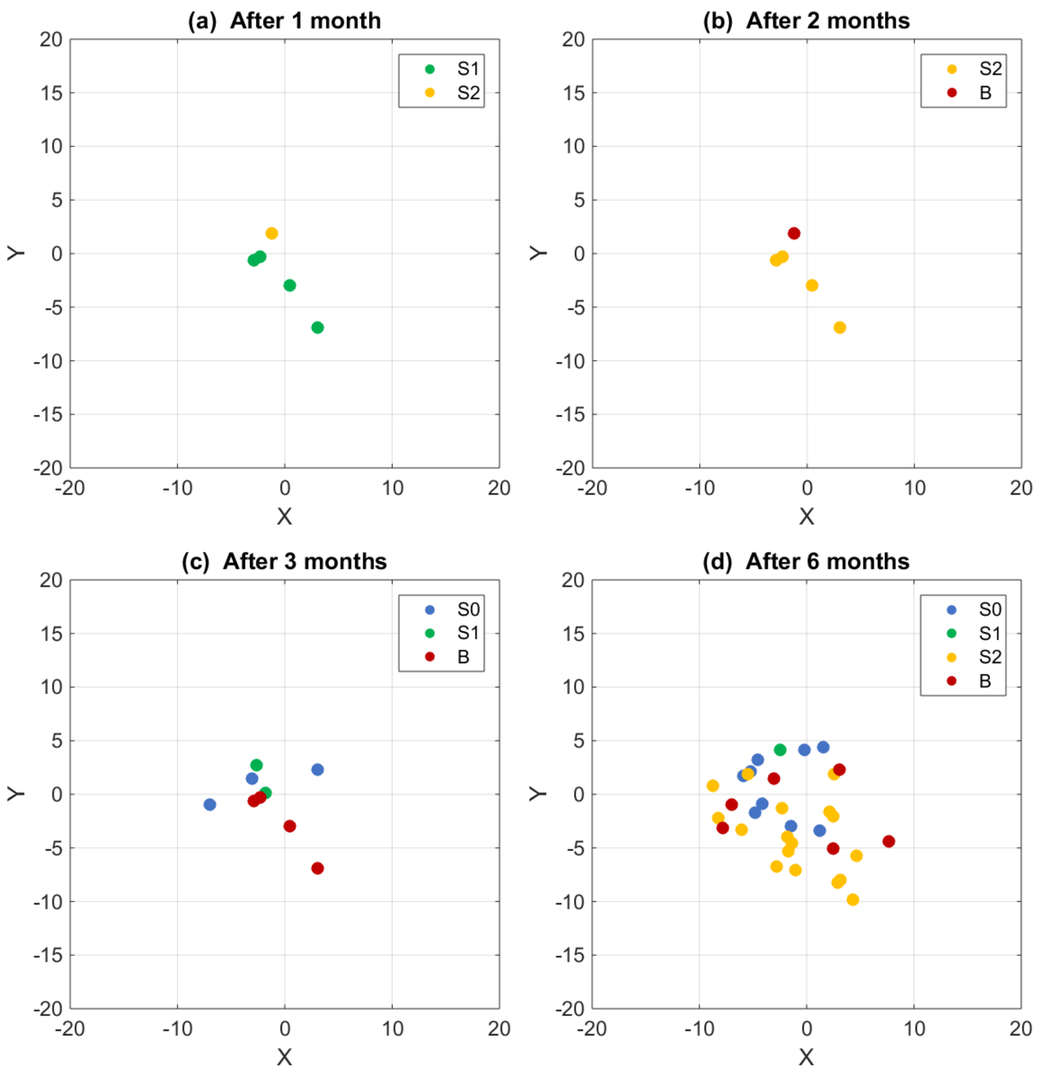
Figure 4 Probability density function of gamma distribution with $\alpha = 3.42,\ \beta = 1$

When the seed is spread by the wind, it goes in various directions, so we assume that the angle Θ follows a uniform distribution between 0° to 360°

$$\Theta \sim U(0, 360)$$

2.3.2 A single experiment of the spatial distribution over a 12-month period

As previously mentioned, we suppose a mature puffball produces $N = 5$ seeds. Each seed has a probability of $P_{short} = 0.25$ to have an accelerated life cycle of three months, and a chance of 75% to have a normal life cycle of four months. We use a computer program to generate random samples for R and Θ using the presumed probability distributions, and calculate the landing position of each seed.



(continued on the next page)

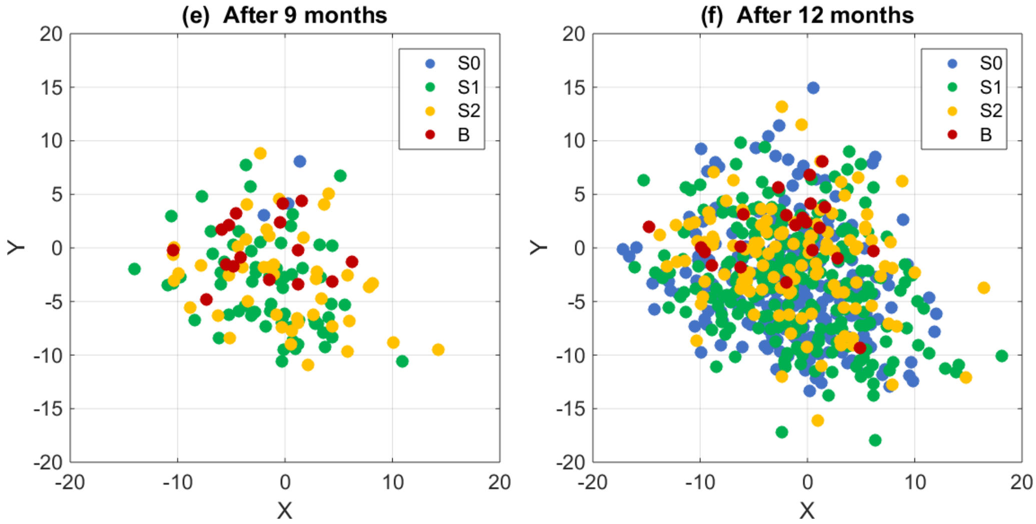


Figure 5 The spatial distribution of the dandelions over a 12-month period

Figure 5(a) shows the distribution of the five seeds produced by the initial seed head. We notice that one of them has shorter life cycle, indicated by yellow color. Figure 5(c) shows the distribution after three months, the dandelion with accelerated life cycle has matured and thus produced its next generation of five new seeds. Three of them have normal life cycle (in blue) and the other two have accelerated life cycle (in green). The four red dots represent the four initial dandelions of normal life cycle which are about to reproduce seeds in the following month. Figure 5(d-f) shows the distribution after 6, 9, 12 months, respectively. We see a significant increase in quantity every 3 months. Over a 12-month period, the initial seed head has produced a total of 633 dandelions, which is consistent with the simulation result using Leslie matrix.

2.3.3 Monte Carlo method for more accurate spatial distribution

Figure 5 only shows the simulation result of a single random experiment. In order to achieve higher accuracy, Monte Carlo method [10] is utilized to repeatedly conduct a large number of experiments. We repeat the same procedure 200,000 times. Figure 6 plots the density of seeds within 20 meters around the initial seed head.

We readily see that the spread of dandelions is omnidirectional, because the density map is rotationally symmetric about the center. This may result from the presumed assumption that the angle Θ follows a uniform distribution, and a new seed is equally likely to be dispersed in every direction.

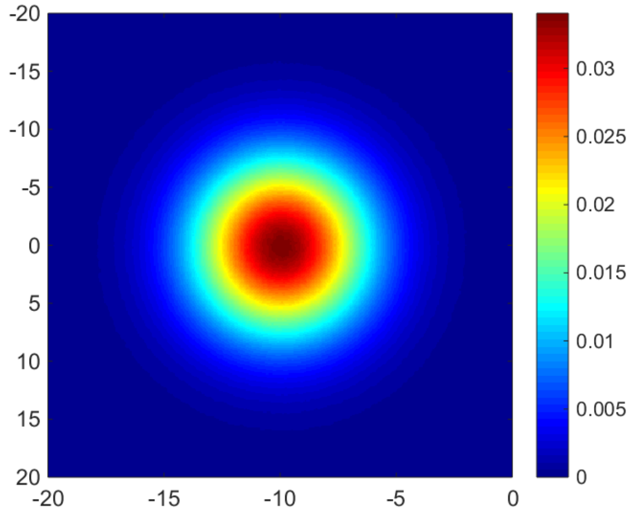


Figure 6 The density map generated by Monte Carlo method

2.4 Estimating the distribution of the dispersal distance after 12 months

As the spread is omnidirectional, we focus on the distance from each dandelion to the center. Figure 7 plots the histogram of the distances.

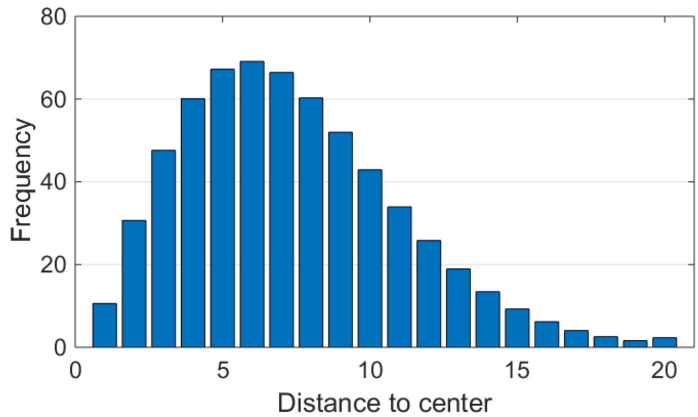


Figure 7 Histogram of the distances of the dandelions to the center

We discover that the shape resembles a gamma distribution, as the number of occurrences rapidly climbs at first and declines slowly after the peak. To test whether the final dispersal distance after 12 months R_{12} follows a gamma distribution, we use χ^2 goodness-of-fit (GOF) test. The null hypothesis H_0 and the alternative hypothesis H_1 are

H_0: the dispersal distances after 12 months follow a gamma distribution $R_{12} \sim \Gamma(\alpha, \beta)$

H_1: the dispersal distances after 12 months do not follow a gamma distribution

To estimate the two parameters of the null distribution $\Gamma(\alpha, \beta)$, we use the following properties.

The expectation (mean value): $E(R_{12}) = \alpha\beta$

The variance: $Var(R_{12}) = \alpha\beta^2$

From our simulation, we calculate the sample mean and the sample variance of the final distances

$$\alpha\beta \approx 6.86 \text{ (sample mean)}$$

$$\alpha\beta^2 \approx 13.34 \text{ (sample variance)}$$

Hence, we estimate the parameters as

$$\alpha \approx 3.52, \qquad \beta \approx 1.95$$

Then, we calculate the expected frequencies f_{exp} as listed in Table 1. The expected frequencies should be no less than 5 to enable a reliable χ^2 GOF test, so we combine the first two categories and the last four categories, yielding a total of 16 categories.

Table 1 Expected and observed frequencies of the final distance from center

Distance	**0-1**	**1-2**	**2-3**	**3-4**	**4-5**	**5-6**	**6-7**	**7-8**	**8-9**	**9-10**
f_{exp}	3.36	22.83	48.73	68.11	76.93	76.49	69.84	60.02	49.29	39.07
f_{obs}	10.57	30.65	47.63	60.08	67.23	69.11	66.44	60.27	51.99	42.90
Distance	**10-11**	**11-12**	**12-13**	**13-14**	**14-15**	**15-16**	**16-17**	**17-18**	**18-19**	**≥ 19**
f_{exp}	30.11	22.67	16.75	12.17	8.72	6.18	4.33	3.00	2.07	4.31
f_{obs}	33.92	25.82	18.94	13.44	9.26	6.20	4.04	2.57	1.60	2.33

*The columns colored in yellow are combined before conducting the χ^2 GOF test.

Compared with the observed frequencies generated by the Monte Carlo method, the test statistic χ^2_{test} is

$$\chi^2_{test} = \sum \frac{\left(f_{exp} - f_{obs}\right)^2}{f_{exp}} \approx 14.33$$

The degree of freedom is

$$df = \text{number of categories} - 1 - \text{number of estimated parameters}$$

$$= 16 - 1 - 2 = 13$$

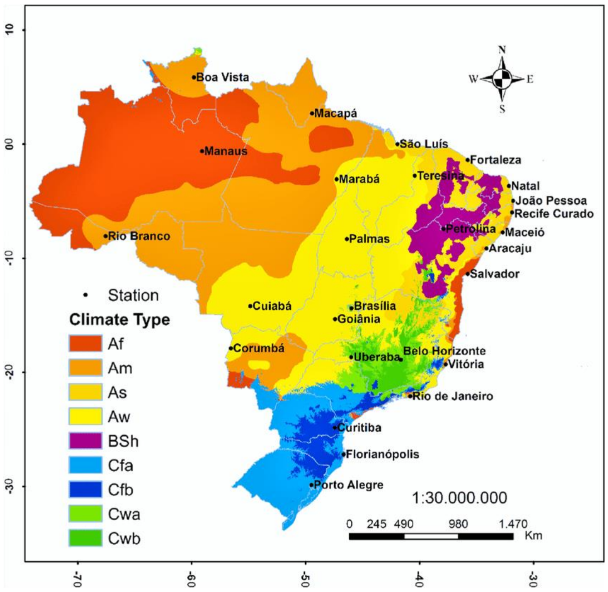
The p-value is

$$p - \text{value} \approx 0.351$$

Since the p-value is greater than the default significant level 5%, we do not have sufficient evidence to reject the null hypothesis H_0. Therefore, the spreading distance of a dandelion after 12 months is considered to follow a gamma distribution $\Gamma(\alpha, \beta)$ with parameter $\alpha \approx 3.53,\ \beta \approx 1.94$. The mean distance is about 6.86 meters.

3 Modeling the spread of dandelions under different weather conditions

To study the effect of different weather conditions on the spread, we select two cities: Rio de Janeiro, Brazil and Mumbai, India. They are chosen from two hemispheres, respectively, and both are in tropical regions where dandelions readily inhabit and spread. The two cities also have quite different climates. Figure 8 shows the climate zones of the two countries, and the locations of the two cities.



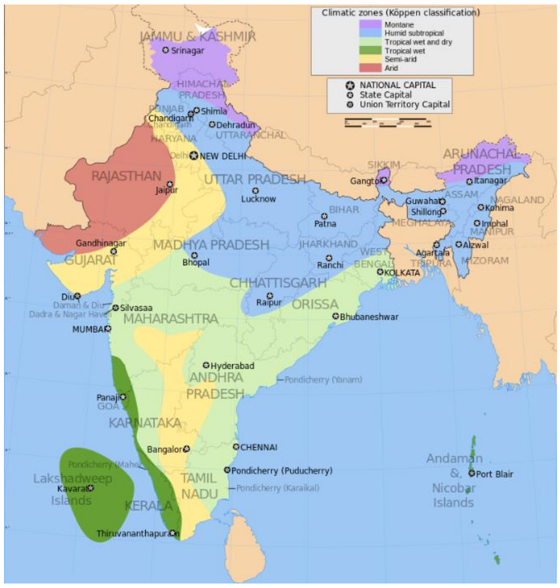


Figure 8 Climate zones of Brazil (left) [11] and India (right) [12]. Rio de Janeiro is located in the southeast of Brazil. Mumbai is located on the west coast of India.

3.1 Climate of Rio de Janeiro and Mumbai

Table 2 Temperature and rainfall by month in Rio de Janeiro [13]

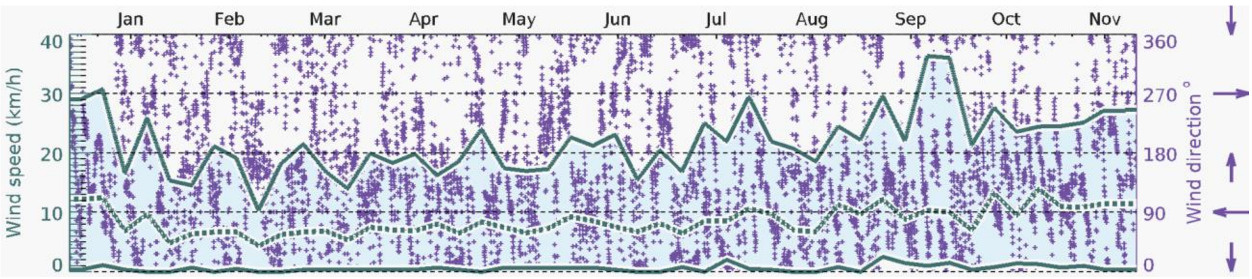
	January	February	March	April	May	June	July	August	September	October	November	December
Avg. Temperature °C (°F)	26.7 °C (80.1) °F	27 °C (80.6) °F	25.9 °C (78.7) °F	24.3 °C (75.8) °F	21.8 °C (71.2) °F	20.8 °C (69.4) °F	20.1 °C (68.3) °F	20.9 °C (69.6) °F	22.2 °C (71.9) °F	23.7 °C (74.7) °F	24.2 °C (75.6) °F	25.8 °C (78.4) °F
Min. Temperature °C (°F)	23.3 °C (73.9) °F	23.3 °C (73.9) °F	22.7 °C (72.9) °F	21.1 °C (70) °F	18.2 °C (64.8) °F	16.8 °C (62.3) °F	16 °C (60.7) °F	16.5 °C (61.6) °F	18.1 °C (64.5) °F	20 °C (68) °F	21 °C (69.8) °F	22.4 °C (72.3) °F
Max. Temperature °C (°F)	31.2 °C (88.2) °F	31.7 °C (89.1) °F	30.2 °C (86.4) °F	28.5 °C (83.4) °F	26.2 °C (79.1) °F	25.8 °C (78.4) °F	25.4 °C (77.7) °F	26.5 °C (79.8) °F	27.5 °C (81.6) °F	28.6 °C (83.5) °F	28.5 °C (83.3) °F	30.1 °C (86.2) °F
Precipitation / Rainfall mm (in)	172 (6)	117 (4)	153 (6)	99 (3)	81 (3)	52 (2)	55 (2)	45 (1)	81 (3)	98 (3)	143 (5)	156 (6)
Humidity(%)	79%	78%	81%	81%	81%	80%	79%	76%	75%	76%	80%	80%
Rainy days (d)	12	10	12	10	9	6	6	6	8	9	12	12
avg. Sun hours (hours)	9.8	10.0	8.8	7.9	7.2	7.0	6.9	7.3	7.2	7.4	7.7	8.8

Table 3 Temperature and rainfall by month in Mumbai [14]

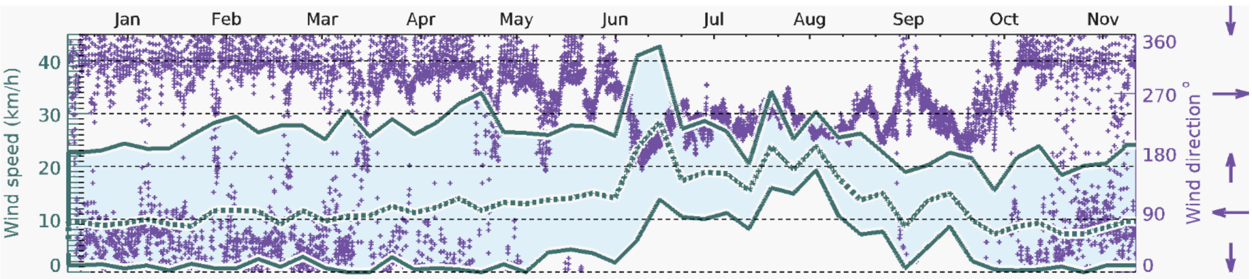
	January	February	March	April	May	June	July	August	September	October	November	December
Avg. Temperature °C (°F)	23.9 °C (75.1) °F	24.7 °C (76.5) °F	26.7 °C (80.1) °F	28.2 °C (82.7) °F	29 °C (84.2) °F	27.4 °C (81.3) °F	25.9 °C (78.6) °F	25.6 °C (78.2) °F	26.1 °C (79) °F	27.3 °C (81.2) °F	27 °C (80.6) °F	25.1 °C (77.2) °F
Min. Temperature °C (°F)	18.8 °C (65.8) °F	19.6 °C (67.2) °F	21.7 °C (71.1) °F	24 °C (75.2) °F	26.2 °C (79.1) °F	25.9 °C (78.6) °F	24.9 °C (76.8) °F	24.5 °C (76.1) °F	24.4 °C (75.9) °F	24 °C (75.1) °F	22.4 °C (72.4) °F	20.1 °C (68.3) °F
Max. Temperature °C (°F)	29.4 °C (84.9) °F	30.2 °C (86.3) °F	32 °C (89.7) °F	32.8 °C (91) °F	32.4 °C (90.4) °F	29.5 °C (85) °F	27.3 °C (81.2) °F	27.2 °C (80.9) °F	28.2 °C (82.8) °F	31 °C (87.8) °F	31.8 °C (89.3) °F	30.5 °C (86.8) °F
Precipitation / Rainfall mm (in)	0 (0)	1 (0)	0 (0)	0 (0)	27 (1)	487 (19)	661 (26)	459 (18)	300 (11)	66 (2)	6 (0)	5 (0)
Humidity(%)	58%	59%	60%	68%	72%	83%	89%	89%	86%	74%	60%	57%
Rainy days (d)	0	0	0	0	3	18	22	22	18	6	1	0
avg. Sun hours (hours)	9.9	10.1	10.5	10.6	9.9	8.8	8.8	9.0	8.5	9.7	10.0	9.8

Table 2 and 3 list the climate data by month of the two cities. The consistently high temperatures above 15°C in both cities provide suitable conditions for the germination and growth of dandelions [15] in the entire year. However, the precipitation conditions are quite different. Influenced by a tropical savanna climate, Rio has abundant rainfall. The rainy season usually lasts from November to March of the next year. In contrast, Mumbai has a distinct wet season from June to September, while the rainfall is scarce in the remaining eight months. In a wet condition, increased humidity changes the pappus shape of the seed, and reduces the dispersal distance [16].

Apart from the temperature and rainfall, the wind condition also has great impact on the spread of dandelion seeds. Figure 9 shows the wind speed and direction of the two cities.



(a) Rio de Janeiro



(b) Mumbai

Figure 9 Wind speed and wind direction in Rio de Janeiro and Mumbai [17]

The wind speed does not vary drastically throughout the year, but the wind directions of the two cities exhibit totally different patterns. The wind direction in Rio de Janeiro is fairly evenly distributed, as the data (in purple) are scattered in the entire range. In Mumbai, however, there is a dominant wind direction in each month. The wind comes from northwest in dry season (from October to the next May), while the wind from west prevails in the rainy season (from June to September). This can be explained by the regional climatic factor that Mumbai is mainly influenced by the wind from the Arabian Sea to its west.

3.2 Estimate the model parameters under different weather conditions

As in Section 2.3, we continue to use a standard gamma distribution, $\Gamma(\alpha, \beta)$ where $\beta = 1$, for the dispersal distance R of a new dandelion seed. In dry weather, the parameter is unchanged, i.e.,

$$\alpha_{\text{dry}} = 3.42$$

The parameter in rainy days is estimated based on the statistic that about 46.1% of the seeds travel less than 1 meter in wet condition [16]. To yield the probability $P(R < 1) = 0.461$, we use a numerical method to obtain the value of α

$$\alpha_{\text{wet}} = 1.41$$

The dispersal direction, represented by the angle Θ, is predominantly affected by the wind direction. For Rio de Janeiro, we assume the angle is uniformly distributed, as in Section 2.3

$$\Theta \sim \mathrm{U}(0,\ 360)$$

To model the prevailing wind direction in Mumbai, we use a normal distribution with the mean aligned with for the dominant direction.

$$\Theta \sim \mathrm{N}(\mu_\Theta,\ \sigma_\Theta^2)$$

From Figure 9(b), in the rainy season from June to September, the wind direction ranges from $160°$ to $320°$. Noting that the coordinate system for Θ is different from the system used in Figure 9(b). The corresponding range for Θ is $-50°$ to $110°$. In addition, we assume it to be the 2σ-interval, $[\mu_\Theta - 2\sigma_\Theta,\ \mu_\Theta + 2\sigma_\Theta]$, where more than 95% of the data lie. Hence, the mean and the standard deviation in wet condition are determined as

$$\mu_{\Theta,\text{wet}} = \frac{110° + (-50°)}{2} = 30°$$

$$\sigma_{\Theta,\text{wet}} = \frac{110° - (-50)°}{4} = 40°$$

In the dry season of Mumbai, the wind direction lies in the interval $[240°,\ 440°]$. The corresponding 2σ-interval for Θ is $[190°,\ 390°]$, and the parameters are therefore calculated as

$$\mu_{\Theta,\text{dry}} = \frac{190°+390°}{2} = 290°$$

$$\sigma_{\Theta,\text{dry}} = \frac{390°-190°}{4} = 50°$$

3.3 Simulating the spread of dandelions

Now we can adjust the probability distribution of the dispersal distance R and the direction Θ in each month. The simulated distributions of the dandelions in Rio de Janeiro after 6, 9 and 12 months are shown in Figure 10.

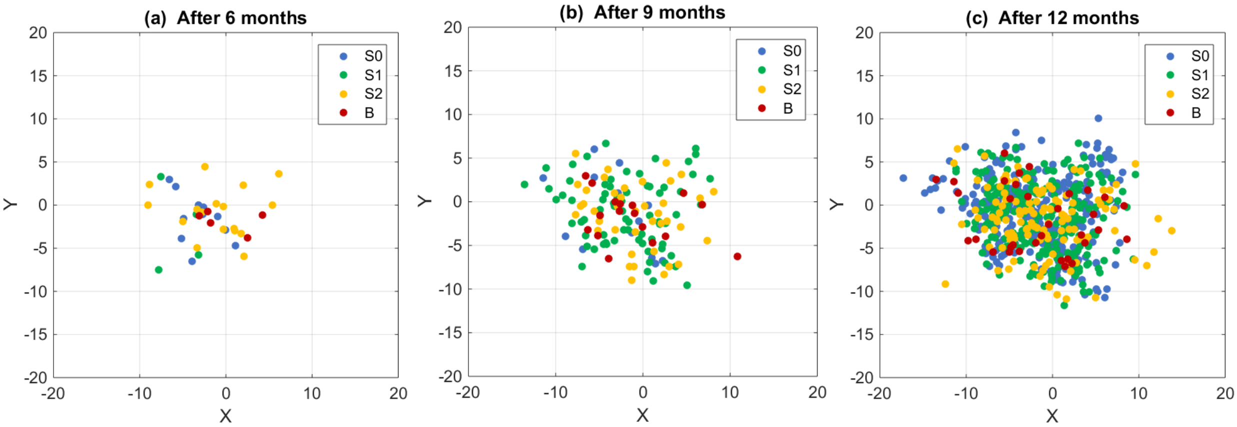


Figure 10 Distributions of dandelions after 6, 9, 12 months in Rio de Janeiro

Table 4 Mean and maximum of the dispersal distances over 12 months in Rio de Janeiro

Month	1	2	3	4	5	6	7	8	9	10	11	12
Mean	1.99	1.99	2.89	4.28	4.28	4.67	5.06	5.59	5.71	5.95	5.90	6.02
Max	4.31	4.31	5.02	9.14	9.14	10.77	12.49	13.70	13.70	15.44	15.44	17.55

Table 4 summarize the mean and maximum distance from center for each month. As expected, the average distance increases over time. Because of the reduced dispersal distance in the rainy season, the final average distance 6.02 meters is slightly lower than the result (6.86 meters) in Section 2.4, and the maximum distance 17.55 meters is much lower than the previous result (25.34 meters). The ultimate center of the dandelions after 12 months is $(1.77, 0.58)$, and it does not deviate much from the initial seed head (at the origin), due to the uniform distribution of the dispersal direction.

The simulated distribution of the dandelions in Mumbai is shown in Figure 11, and the average and maximum dispersal distance are listed in Table 5.

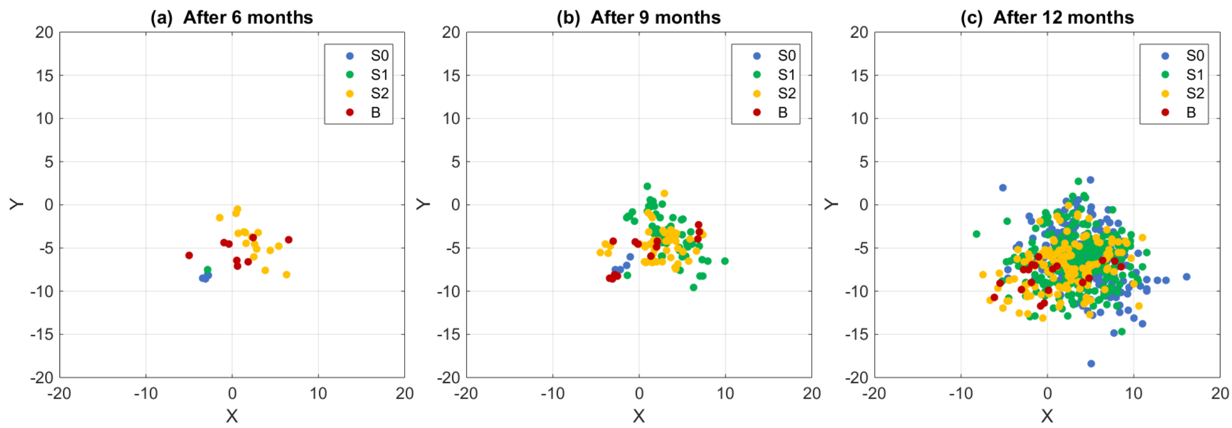


Figure 11 Distribution of dandelions after 6, 9, 12 months in Mumbai

Table 5 Mean and maximum of the dispersal distances over 12 months in Mumbai

Month	1	2	3	4	5	6	7	8	9	10	11	12
Mean	2.86	2.86	5.08	5.49	5.49	5.91	6.10	5.64	5.69	6.94	7.97	8.14
Max	5.24	5.24	9.47	10.21	10.21	10.21	10.21	11.86	11.86	13.61	15.76	19.09

After 12 months, both the mean distance and the maximum distance are higher than the respective statistic obtained in the previous section, largely due to the consistent western wind throughout the year. Driven by the wind, the final center of the dandelions $(3.32, -6.50)$ is to the southeast of the initial seed head.

4 Impact factor for identifying invasive species

In previous sections, we see that a single matured dandelion is capable of producing numerous developed seedlings within a year. Ecological reports also state that dandelions have aggressive root system and thrive on disturbed soil. Therefore, dandelion is usually labelled as a weed or an invasive plant. In this section, we attempt to develop a quantitative approach to evaluate the negative impacts of a plant on the environment. The idea is motivated by the research [18], which proposes 20 questions addressing a species' negative impact on biodiversity. Over 100 people with expertise in biology and management are devoted to evaluating about 3500 nonnative plant species and create a national list prioritized by their negative impact [18].

We choose to focus on the following seven factors in the evaluation. Some factors address the impact on other species and the entire ecosystem, while other factors are more related to the current distribution and trends in distribution and abundance. For each factor, an empirical weight is assigned in the form of points out of 100, so that its relative importance can be taken into account. Furthermore, several levels together with descriptors are designed for rating. The total points achieved from the seven factors is used as an impact factor that assesses the extent that the plant is invasive and threatening to the environment. The details of the factors are introduced below.

4.1 Assessment criteria

Criterion A: Impact on ecosystem (maximum 30 points)

Point	Description
0	The plant has negligible negative impact.
5	The plant has very limited impact on the ecosystem, regarded as a normal plant.
10	The plant has small impact on the ecosystem, considered a weed or disturbance.
15	The plant has noticeable impact on the ecosystem, considered to have environmental risk.
20	The plant has moderate impact on the ecosystem. It might lead to significant reduction of some native species in the system.
25	The plant has moderate to high impact on the ecosystem. It evidently disturbs the agriculture.
30	The plant significantly alters the ecosystem process and structure.

Criterion B: Impact on plants (maximum 10 points)

0	The plant has negligible harm to native plants.
2.5	The plant has limited impact on plants. It occupies a small proportion of land originally occupied by native plants.
5	The plant has moderate impact on plants. It occupies a moderate amount of land originally occupied by native plants, and competes with native plants for resources.
7.5	The plant has moderate to high impact on plants. It causes significant decease in resources available for native plants.
10	The plant brings severe threat and causes endangered native plant(s).

Criterion C: Impact on animals (maximum 5 points)

0	The plant has negligible impact on the food chain and animals.
2.5	The plant has moderate impact on the food chain and animals.
5	The plant completely alters the food chain, and threatens the animal(s) in the same region.

Criterion D: Range of impact (maximum 20 points)

0	The plant has limited range of impact, typically within a range of 1 meter.
5	The plant has small range of impact, typically around 5 meters.
10	The plant has moderate range of impact, typically around 10 meters.
15	The plant has large range of impact, typically around 50m
20	The plant has maximum range of impact, typically more than 100m

Criterion E: Reproduction (maximum 20 points)

0	The plant has very slow growth. It usually flowers about 9 months after a seed is planted.
5	The plant has slow growth. It develops slowly, and flowers in about half a year after a seed is planted.
10	The plant has moderate growth speed. It flowers in about 4 months after a seed is planted. It thrives on disturbed environment.
15	The plant has high growth speed. It flowers in about 2-3 months after a seed is planted. It thrives on disturbed environment.
20	The plant has extremely high growth speed. It flowers within about 1 month after a seed is planted. It thrives on disturbed environment.

Criterion F: Potential of spread (maximum 5 points)

0	The plant is difficult to spread in nature without human intervention.
2.5	The seeds can be distributed in multiple ways. The plant has moderate spread distance.
5	The seeds can be distributed by birds, winds, or unknowingly humans and travel great distances.

Criterion G: Difficulty in management (maximum 10 points)

0	The plant can be easily removed completely by pesticides or other chemicals
2.5	The plant can be easily controlled.
5	It is moderately difficulty to control the growth and the range of the plant.
7.5	It is difficult and costly to control the growth and the range of the plant.
10	It is extremely difficult and resource-demanding to control the spread of the plant.

The points in all the 7 factors are added up to obtain the impact factor, to evaluate the overall negative impact of the plant. We divide the total mark into five categories, as follows.

Table 6 Five categories based on the impact factor

Total mark	Description
0-20	The plant has limited impact on the native species and ecological community.
21-40	The plant has small impact on the native species and ecological community.
41-60	The plant has moderate impact on the environment, and disruption to native species.
61-80	The plant has high impact on the environment, considered a threat to the ecosystem.
81-100	The plant has extremely high negative impact on the environment, considered as a great threat to the ecosystem.

4.2 Testing the impact factor on invasive plants

4.2.1 Impact factor for dandelion

When developed in an ecosystem, dandelions have a relatively low ecological impact and do not cause substantial damage to ecosystems. The major nuisance is their persistence in yards and managed spaces where growth is unwanted. Often considered an annoying weed, dandelions can inhabit highly disturbed environments like lawns [1]. Dandelions exhibit allelopathic traits that allow them to outcompete native plants. Their pollen contains allelochemicals that reduce native species' reproductive success when transferred. Their allelopathic root exudates also hinder the growth of neighboring plants [1]. Dandelion pollen is detrimental to some native bees. When dandelion pollen was fed to solitary bees, larval development was found to be stunted or incomplete. Native pollinators are better served by the native plants that bloom in spring, including trees and wildflowers [2]. Based on the above information, we rate 10, 7.5, 5 for criteria A, B, C, respectively.

It usually takes 85 to 95 days for a dandelion seed to blossom, and another two weeks to mature into a seed head [1]. Dandelion seeds are easily detached from the seed head in a gust of wind, and capable of travelling long distances. 99.5 percent of dandelion seeds land within 10 meters, 0.05 percent travel more than 100 meters, and 0.01 percent would travel greater than 1km [7]. It could take several seasons to fully eradicate dandelions in a yard. Dandelion has a great capacity for resprouting from damaged roots. Root fragments as small as 0.05 inch in diameter by 0.25 inch in length can establish new plants [1]. Accordingly, we rate 10, 15, 5, 7.5 for criteria D to G, respectively.

Finally, the total points for dandelion is

$$\mathrm{IF}_{dandelion} = 10 + 7.5 + 5 + 10 + 15 + 5 + 7.5 = 60 \text{ points}$$

which indicates that dandelions have a moderate impact on the ecological community. It is consistent with the consensus that dandelion is considered a weed, and has certain disruption to native species.

4.2.2 Impact factor for two other invasive plants: Creeping Charlie and English ivy

Glechoma hederacea, known as Creeping Charlie, is recognized as an invasive species. It was introduced to North America from Europe in the 17th century. Creeping Charlie spreads or creeps across the top of the ground via surface roots or runners, called stolons, and it has an aggressive rhizome root system that spreads horizontally below the ground. The plant can propagate through seeds and plant fragments. It readily invades areas where other plants fail to thrive, such as shady areas or soil with limited moist and fertility. Once established, it creeps to sunny sites, crowding out native grasses and plants. Its rapid colonization displaces less robust flora, demonstrating substantial invasive impact on local ecosystems [19]. There is limited evidence that Creeping Charlie poses harm to animal populations. They are not toxic if not consumed in huge amounts. Based on the above information, we rate the plant 25, 7.5, 0, 20 for criteria A to D, respectively.

Creeping Charlie has four seeds on each flower, and germinate within a few days after landing on moist soil. It is well-known for its rapid growth once it stops flowering [19]. The plant's extensive creeping system makes it difficult to eradicate by hand-pulling. A dethatching tool might cause damage to the lawn or spread fragments which undesirably induce development of new plants. Creeping Charlie can be killed by borax, but it is toxic to animals like ants, and causes adverse long-term effects on soil and groundwater. Accordingly, Creeping Charlie gets 15, 5, 20 points for criteria E to G, respectively.

In total, the impact factor for Creeping Charlie is

$$\mathrm{IF}_{cc} = 25 + 7.5 + 0 + 20 + 15 + 5 + 20 = 92.5$$

which shows that the plant is a great threat to the ecosystem.

Another recognized invasive species to be tested is Hedera helix, also known as English ivy. It has invaded California and the northwestern United States, and becomes an issue in regions near the coast [20]. English ivy weakens and kills trees by engulfing branches and blocking sunlight from the tree leaves, preventing them from making food that fuels the tree's growth and ensures its viability. Eventually, trees can become completely smothered or die from progressive weakening. As ivy climbs higher in a tree, branch dieback advances from lower to higher branches, and make infested trees more likely to topple in extreme weather situations. English ivy also endangers whole ecosystems by forming dense and extensive monocultures that displace native plants. The leaves and berries of English ivy contain a toxin. Clinical signs drooling, vomiting, and diarrhea are expected following ingestion by pets. They are also toxic to many pets, including dogs, cats, and horses [21]. Therefore, we rate English ivy 30, 10, 5 for criteria A to C.

English ivy grow swiftly. It can reach 90 feet and spread vegatatively outward through its long vines and climb over almost any obstacle. English ivy also reproduces from seeds that are dispersed by birds when they eat the English ivy fruit. Removal of English ivy is fairly straightforward, with minimal follow-up if the area is removed successfully the first time. However, removal can be time consuming in high density areas. Therefore, English ivy achieves 20, 5, 15, 5 points for criteria D to G, respectively.

The impact factor for English ivy is $\mathrm{IF}_{ivy} = 90$ points, suggesting a serious threat to the ecosystem.

4.3 Testing the impact factor on non-invasive plants

Apart from invasive plants, our impact factor is also tested on non-invasive species, to see whether the scale is valid to distinguish the two categories.

Watermelons do not appear on any state or national list of invasive species. They grow to full potential even on land that has minimal nutrients and fertilizer to grow. Watermelons only take 235 liters of water to produce a kilogram of fruit [23]. They need 65 to 100 days to grow. They have very limited impact on other plants and they are safe to eat. Though watermelons can spread out and loop over fences, they can be easily controlled. Therefore, only 5 points are awarded for criteria D and F, respectively, yielding an impact factor of $\text{IF}_{watermelon} = 10$ points.

Tulips are also known to be non-invasive. Rather than harming the ecosystem, tulips provide benefits. As an important resource for pollinators, tulips offer foraging opportunities, especially for bees. Their flowers produce both pollen and nectar, providing energy to pollinators while enabling the plant's sexual reproduction. There is also no evidence suggesting tulips negatively impact other species or animal populations. Tulips require an extensive flowering period, ranging from 15 to 30 weeks [23]. Tulip blooms produce seed pods that can be spread by the wind, rain, or animal activity. Consequently, only 5 point is assigned for the potential of spread, leading to an impact factor of $\text{IF}_{tulip} = 5$.

Compared with the impact factors of dandelion, Creeping Charlie and English ivy, those for non-invasive species are significantly lower. The distinctiveness of impact factors for the two categories enables a clear identification and classification.

5 Conclusion and discussion

In this paper, several mathematical models are proposed to characterize the spread of dandelions. The discrete-time model using Leslie matrix is capable of estimating the growth of dandelions in different stages and the interaction between them. The calculation is simple because the ecological features of the plant are integrated into a transition matrix, and populations of the stages are concatenated into a column vector.

Spatial distribution over time is simulated by a Monte Carlo method with presumed probability distribution for the key random variables. Gamma distribution is suitable for the dispersal distance, as it only generates positive samples and has a long tail to model the great travel distance in rare cases. The distribution for the wind should be selected based on the regional climatic data. Normal distribution is recommended if a predominant wind direction is evident, while uniform distribution is more appropriate for more random wind directions. Our simulation results show that the dispersal distance of the dandelions reproduced in a year is likely to follow a gamma distribution, if wind direction is uniformly distributed. Long rainy season would hinder the spread whereas consistent wind direction facilitates long-distance propagation.

In addition, we also design an impact factor for quantitatively evaluating how invasive a species is to the environment. Multiple factors are included to assess the overall negative impact of a plant in a more comprehensive manner. A set of ranks with descriptors are proposed so that the rating would be more consistent across different examiners. The impact factor for the dandelion indicates the plant has moderate impact on the ecosystem and the environment. In addition, comparison of the impact factors evaluated for invasive and non-invasive species shows that the metric is highly distinctive for the two classes.

One major advantage of the methodology is that it offers a set of model parameters, which can be customized in other scenarios of similar academic purposes. Expertise in certain areas can also help to determine more accurate parameters. The methodology is scalable and adaptable, and able to be extended to a wider range of applications.

6 References

[1] Stewart-Wade, S. M., S. Neumann, L. L. Collins, and G. J. Boland. (2002). The biology of Canadian weeds. 117. Taraxacum officinale G. H. Weber ex Wiggers. Canadian Journal of Plant Science 82:825-853.

[2] “Find out more about invasive species” (https://www.invasive-species.org/about/) Retrieved 2023-11-01

[3] “Taraxacum” (https://en.wikipedia.org/wiki/Taraxacum). Wikipedia. Retrieved 2023-11-01

[4] “Dandelion” (https://www.agric.gov.ab.ca/app107/). Ministry of agriculture and irrigation. Retrieved 2023-11-01

[5] “Dandelion” (https://www.canr.msu.edu/weeds/extension/dandelion). Department of Plant, Soil and Microbial Sciences, Michigan State University. Retrieved 2023-11-01

[6] Honek, A., Z. Martinkova, and P. Saska. (2005). Post-dispersal predation of Taraxacum officinale (dandelion) seed. Journal of Ecology 93:345-352.

[7] Tackenberg, O., P. Poschlod, and S. Kahmen. (2003). Dandelion seed dispersal: The horizontal wind speed does not matter for long-distance dispersal - it is updraft! Plant Biology 5:451-454.

[8] Leslie, P.H. (1945). The use of matrices in certain population mathematics. Biometrika, 33(3), 183–212.

[9] Devore, J. L. (1995). Probability and Statistics for Engineering and the Sciences (Vol. 5). Belmont: Duxbury Press.

[10] Metropolis, N., & Ulam, S. (1949). The monte carlo method. Journal of the American statistical association, 44(247), 335-341.

[11] Jerszurki, Daniela & Souza, Jorge & Silva, Lucas. (2019). Sensitivity of ASCE-Penman–Monteith reference evapotranspiration under different climate types in Brazil. Climate Dynamics. 53. 10.1007/s00382-019-04619-1.

[12] “Climate of India”, (https://en.wikipedia.org/wiki/Climate_of_India), Wikipedia, Retrieved 2023-11-06

[13] “Rio de Janeiro climate”, (https://en.climate-data.org/south-america/brazil/rio-de-janeiro/rio-de-janeiro-853/), Climate Data, Retrieved 2023-11-06

[14] “Mumbai climate”, https://en.climate-data.org/asia/india/maharashtra/mumbai-29/), Climate Data, Retrieved 2023-11-06

[15] Luo, J., and J. Cardina. 2012. Germination patterns and implications for invasiveness in three Taraxacum (Asteraceae) species. Weed Research 52:112-121.

[16] Seale, M., Zhdanov, O., Cummins, C., Kroll, E., Blatt, M., Zare-Behtash, H., Busse, A., Mastropaolo, E., Viola, I.M. and Nakayama, N., 2019. Moisture-dependent morphing tunes the dispersal of dandelion diaspores. Available at SSRN 3334428.

[17] “Weather Archive Rio de Janeiro” & “Weather Archive Mumbai”, （https://www.meteoblue.com/, Retrieved 2023-11-10

[18] Randall, J.M., Morse, L.E., Benton, N., Hiebert, R., Lu, S. and Killeffer, T., 2008. The invasive species assessment protocol: a tool for creating regional and national lists of invasive nonnative plants that negatively impact biodiversity. Invasive Plant Science and Management, 1(1), pp.36-49.

[19] “Glechoma hederacea” (https://en.wikipedia.org/wiki/Glechoma_hederacea), Retrieved 2023-11-10

[20] “Hedera helix” (https://en.wikipedia.org/wiki/Hedera_helix) , Retrieved 2023-11-10

[21] “Toxicity to pets” (https://www.petpoisonhelpline.com/poison/english-ivy/), Retrieved 2023-11-10

[22] “Watermelon Benefits and Side Effects” (https://www.healabel.com/watermelon-benefits/), Retrieved 2023-11-11

[23] “How Long Do Tulips Bloom” (https://meadowlarkjournal.com/blog/how-long-do-tulips-bloom #google_vignette) , Retrieved 2023-11-11