Procedural modeling of plant ecosystems maximizing vegetation cover

Cristina Gasch1 · José Martínez Sotoca1 · Miguel Chover1 · Inmaculada Remolar1 · Cristina Rebollo1

Received: 12 April 2021 / Revised: 29 July 2021 / Accepted: 3 January 2022 / Published online: 2 March 2022 © The Author(s) 2022

Abstract
Vegetation plays a major role in the realistic display of outdoor scenes. However, manual plant placement can be tedious. For this reason this paper presents a new proposal in the field of procedural modeling of natural scenes. This method creates plant ecosystems that maximizes the covered space by optimizing an objective function subject to a series of constraints defined by a system of inequalities. This system includes the constraints of the environment taking into account characteristics of the terrain and the plant species involved. Once the inequality system has been defined, a solution will be obtained that tries to maximize the radius of the projected area of the trees and therefore the extension of the vegetation cover on the ground. The technique eliminates the trees that do not achieve a minimum growth radius, simulating the typical competitive process of nature. Results show the good performance and the high visual quality of the ecosystems obtained by the proposed technique. The use of this kind of optimization techniques could be used to solve other procedural modeling problems in other fields of application.

Keywords Optimization · Inequality systems · Plant ecosystem · Procedural distribution · Botanical constraints · Visual Realism

Inmaculada Remolar
remolar@uji.es
Cristina Gasch
cgasch@uji.es
José Martínez Sotoca
sotoca@uji.es
Miguel Chover
chover@uji.es
Cristina Rebollo
rebollo@uji.es

1 Institute of New Imaging Technologies, Universitat Jaume I, 12006 Castellón, Spain
1 Introduction

The representation of outdoor scenes is currently a very popular field of research due to the huge number of applications that require them, such as ecosystem simulation [20, 26], video games [27], or geographic information systems [6, 39]. Procedural modeling of individual plants and trees has been widely studied in the literature [38]. Nevertheless, distributing them automatically so as to fill the space available is still an open problem. Smelik et al. [41] reviewed some studies that use procedural techniques for the creation of natural environments and virtual worlds. Specifically, they estimate the virtual position of the plants from a procedural calculation of their distribution. They conclude, however, that these techniques achieve unrealistic results, since they lack the ability to translate ecological input data into procedural distributions.

In order to increase the realism of the outdoor scenes, it is necessary to take into account some botanical information in the procedural placement, such as the biotic and abiotic characteristics of the species, as well as the volume that they take up [11, 12, 36]. Smelik et al. [41] stated that an ecological plant distribution with sufficient density and variety, where each vegetation species is placed in a predetermined position, is a difficult challenge to achieve [2, 3, 8, 14].

According to Lane and Prusinkiewicz [29], ecological modeling techniques can be divided into two groups: local-to-global and global-to-local. On the one hand, the procedural methods classified as local-to-global treat plants as individual entities to which rules are applied in order to simulate factors such as reproduction, growth, competition, interaction, variation or adaptation to the environment [15]. This biological model of plants predicts patterns of vegetation distribution in an ecosystem where plants grow, adapt to the environment, reproduce and compete for space. These methods require a high consumption of resources and the processing time for each simulation cycle is also high. On the other hand, the procedural methods classified as global-to-local calculate the positions of the plants directly from a globally defined environment. Thus, they include aspects related to the terrain such as height and slope, climate characteristics or vegetation cover density.

In order to avoid the problems of local-to-global techniques [29], this work presents a method for the distribution of trees and plants in natural environments through the use of linear programming techniques. In this case, the objective function to be maximized is the projected area of the set of plant species. This means trying to achieve the largest possible vegetation cover with certain constraints. These constraints are defined by a system of inequalities that includes the biological characteristics of the plant species, such as the maximum growth radius or the relationships between the growth radii of the different species.

In this sense, after making an initial irregular distribution of the seed positions where the different plant species can be placed, the user is allowed to indicate some initial growth regions of each of the plant species. Finally, a distribution of the species is made taking into account for each placement position both the characteristics of the plants, as well as the characteristics of inclination and height of the terrain. Once the final distribution of the plant species has been determined, the system of inequalities that includes the radii of each of the plants as variables is created.

The novelty of this methodology is based on providing a global solution to the distribution of plants and trees, avoiding the drawbacks of iterative techniques. Furthermore, the technique can easily be adapted to the definition of other types of objective functions and any modification or inclusion of new constraints. Finally, it should be noted that the
implementation of the system is simple and robust, which makes it ideal for solving this type of problem.

The paper is organized as follows. Section 2 considers the state of the art related to the existing techniques in the procedural generation of natural environments. Section 3 briefly describes the method operation scheme. Section 4 analyzes the distribution of seed positions, and Section 5 describes the labeling process for each seed positioned on the terrain. Section 6 describes the space-filling process performed by the inequality systems, which makes it possible to obtain a high degree of realism. Section 7 analyzes the results obtained by testing the method and, lastly, Section 8 presents some conclusions and observations related to future work.

2 Previous work

Designing a realistic outdoor scenario with an ecological distribution of plants is not an easy task to achieve. The biotic and abiotic characteristics of the vegetation species \[10, 44\] are essential in the process of obtaining the exact positions of the plants in a terrain where a high density and variety of species are required \[32\]. Onrust et al. \[34\] classified the techniques for plant species placement in two main groups: those based on ecological modeling techniques and those based on the procedural generation of ecosystems.

Studies classified in the first group distribute the plants based on real data related to the terrain or the vegetation species to be represented. These techniques can, in turn, be classified in two sub-groups depending on the source of the data used as input. On the one hand, dynamic ecological modeling techniques simulate ecological processes taking into account data that can be extracted from a raster map, such as height, biomass or vegetation cover \[6, 20, 40, 42\]. On the other hand, neutral ecological modeling techniques \[30\] generate a distribution of plants on a terrain following a grid structure that is based on some vegetation species information.

According to Onrust et al. \[34\], the second group is related to the procedural generation techniques that perform an automatic distribution of plants and trees. Studies in this group estimate their position using different algorithms. Nevertheless, these techniques usually fail to create a distribution of plants with sufficient variety and realism because no ecological input data are considered. Over the years, several studies that address this problem have appeared.

Deussen et al. \[15\] proposed the framework of the Open L-Systems model \[31\] to simulate the individual-based ecosystem models. The growth of an individual ecosystem is represented using the dynamic behavior model of Firbank et al. \[18\]. This model starts from a set of randomly distributed circles with initially random radii that represent plants growing in the ground. In a second phase, control by the user is incorporated through a gray-scale image that represents the density of the plants. Individual plant positions that fit these densities are created using the Floyd-Steinberg error diffusion algorithm \[19\]. Finally, the points produced by this algorithm are altered slightly to make the distribution more random. Nevertheless, the resulting distribution of the plants is too uniform. Moreover, the cost of the process in terms of time is high, according to Lane et al. \[29\].

These authors \[29\] proposed the symbiosis of multiple vegetation species to generate plant distributions. They combine a probability value with a dart-throwing algorithm that allows plants to be placed in their most probable areas. In addition, each plant can produce an inhibition or expansion effect on the remaining plants, by updating the probability field.
around it. However, the grammatical elaboration method turns out to be quite complex and the computation times are relatively high.

A method called FON (Field Of Neighborhood) was proposed by Bauer et al. [7]. It describes a circular zone of influence around the plant, the radius of which determines the distance from which it can interact with its neighbors. This radius depends on the soil characteristics, size, and nutritional needs of the plant and the space needed for that nutrition. Based on FON, Alsweis et al. [4] proposed a method that simulates plant communities of two different species in evolution. Later, in [5], they calculated the distribution of the plants by combining Poisson Disk Distributions (PDDs) [28] with Wang-tiles [17], although neither the distribution by species nor the interactions between them are considered.

Weier et al. [43] extended the previous method and generated large areas of vegetation using Wang-tiles with reusable distributions that allow a terrain to be covered. Each Wang-tile is generated using PDDs with a variable radius. The tiles are organized in a multilevel structure and treated with a nested KD-tree, which allows them to be reused during visualization. In addition, they propose a solution to represent plants whose area of influence is on the edges of the tiles. The main problem with this proposal is that the size of the disks that line the ground tend to be too homogeneous.

Onrust et al. [33] presented a procedural plant placement technique that combines the Poisson Disk Distribution with Wang-tiles, thereby establishing an ecological modeling technique. Results were only tested with isolated plants that had problems with the position of the impostors with respect to the camera. The work proposed by Onrust et al. [34] is based on vegetation location information resulting from the combination of several parameters such as the Normalized Difference Vegetation Index (NDVI), height-map, biomass or coverage maps. Gasch et al. [23] presented a method of automatic vegetation placement that takes as its input the height-map of a terrain where a route is represented [24]. The method extracts the values necessary to distribute the vegetation species from the height-map and combines them with a noise distribution based on Perlin noise [35]. By doing so, it becomes possible to obtain placement patterns with a greater probability of occurrence in favorable points on the map and near the route. However, this method does not take into account the development and growth of plants.

Li et al. [30] proposed a procedural model based on graphs that generates realistic vegetation distributions with different compositions and configurations adjusted to the metric of a specific landscape created from patches. The metrics used include the number of tree species, their proportions in the landscape, the number of patches that have to be populated by each species and their average shape. The seeds are placed randomly and their species are assigned in order to obtain the final distributions of the trees. The method establishes the distribution of each species with a time efficiency that satisfies the requirements to build large-scale virtual forests.

A more generic method to describe growth and simulation of ecosystems is presented in Growth-Grammars [16], which provides the user with access to a tool for developing growth grammars through an interactive modeling platform, GroIMP. This grammar is executed through programs written in XL (a Java-based growth grammar language) and visualized by XL4C4D (an XL-plugin for CINEMA 4D). Although this method allows populations of plants to be generated in small areas, it does not reuse certain components to cover more extensive areas.

Benes et al. [9] described a method to distribute vegetation in urban environments procedurally. They addressed the problem of spatial distribution of vegetation by combining a user-guided simulation system with an interactive procedural system that integrates plants
into the urban environment. Starting from the 3D geometry of an urban layout with initial conditions and parameters, procedural rules are inferred. A manageability level is calculated for each area of the urban space, which defines the degree of influence between the simulation of the wild ecosystem, where plants compete freely for resources and seeds, and the ecosystem is managed by the user and the plants grow only under well-defined conditions. The method presented takes some of the weaknesses found in the literature into account and proposes a solution that includes most of the detected shortcomings. How it works is analyzed in the following sections.

3 Plant covering method

The main objective of this method is to maximize the size of the vegetation cover in nature scenes. To obtain a realistic representation, the method considers three main processes: the distribution of positions where seeds can be placed, the assignment or labeling of the seeds to those positions and the resolution of the system of inequalities that will allow to maximize the area of vegetation cover (see Fig. 1).

The initial distribution of the seed positions can be done with any strategy. In this case, a regular distribution has been made whose positions are later slightly modified to obtain a more natural irregular appearance.

Once the positions where the different plant species can grow have been determined, a subset of the positions is chosen and an initial labeling is carried out taking into account some characteristics of the terrain, such as slope and height. In this way, an initial distribution of the plants is already obtained. Subsequently, the final labeling of all the seed placement positions is carried out taking into account the closest neighbors.

The method ends with the creation of a system of inequalities that poses the constraints regarding maximum radius and distance relationships between species. The resolution of the system will determine the radius of each of the trees or plants that maximizes the area of vegetation cover. After obtaining these radii, all the species that have not been able to reach their minimum radius are eliminated and the system of inequalities is redone a second time to obtain in this case a more precise solution.

4 Seed dissemination process

The method presented here starts by determining the areas without vegetation on the terrain. A criterion to choose these areas can be, for example, zones situated higher than a

![Fig. 1 Scheme of the proposed model](image_url)
certain height, or areas where some paths, lakes or rivers are located. However, other ways to achieve these non-vegetation areas can be considered in the case of computer-generated terrains.

In this work, the terrain was obtained using the Perlin noise method [35]. The process starts by assigning a value that represents the possibility of being populated by plants (between 0 and 1) to each point in the terrain that has been generated. A threshold is then established to determine the areas that can simulate grasslands, sandy areas, rocky areas, etc.

The initial positioning follows a regular grid distribution where the seeds are separated by a distance of $d_{\text{spacing}} = 5$ meters. For this purpose, a relation of 1 meter = 1 pixel was considered in its representation within the placement map. Then, in order to break up the regular location of the seed positions, the method performs a deviation of the seed positions $s_p$, by modifying their $(x, y)$ coordinates with a random value using a uniform distribution of random number generator. This random value will always be less than half the spacing used in the generation of the grid, in order to prevent two elements from overlapping.

The final result of the redistribution can be seen in Fig. 2a, where areas with no vegetation are represented in black, and a closer view can be observed in Fig. 2b. Both figures show the deviation suffered by the seed positions, thereby eliminating the unrealistic regular appearance and covering all the terrain.

5 Seed labeling process

The next process is to determine the species to which each seed belongs. Initially, only a sub-set of the distributed seeds are evaluated by taking into account several abiotic features of the terrain and a simulation of the grouping of plants of the same species that happens in nature. Finally, the influence of the neighboring vegetable species determines the label of the rest of the seeds that remain unlabeled on the terrain.

Only the height and the slope of the terrain have been considered in this work. These abiotic characteristics have been included in a linear equation that establishes the appropriate vegetable species depending on the features taken into account. However, more features could easily be added to the method.

![Fig. 2](image-url)
5.1 Influence of the abiotic characteristics

The slope and the altitude of the seeds are data available from the height-map. Both characteristics are taken into consideration to analyze the possibility of each vegetable species growing at that location. In their work, Galin et al. [21] stated that the ability of the species to live is limited by multiple conditions, so that although a vegetable species has a perfect soil, it cannot grow if it does not have sufficient temperature or light. In this work, if the position of a seed is not in the range of the altitude or slope values where the vegetable species under consideration can grow, it will not be able to develop and it will die.

Let $V$ be the set of vegetable species to be distributed and $v$ a particular species in that set. Since $s_p$ is the position of the seed, the altitude of the seed can be obtained directly in meters from the height-map $\text{height}(s_p)$ (Fig. 3a). The slope where a seed is situated, $s_p$, is computed using the following equation:

$$
slope(s_p) = \arccos \left( \frac{N \cdot N(s_p)}{|N| \cdot |N(s_p)|} \right)
$$

(1)

where $N$ is the vector normal to the horizontal plane and $N(s_p)$ is the normal vector at the point where the seed is situated. The $\text{slope}(s_p)$ is measured in degrees. Figure 3b shows the slope characteristics of the terrain under consideration.

Analyzing the plant distribution depending on the height, the distribution of plants per species $w(s_p, v)$ is given by the following equation:

$$
\forall v \in V \\
w(s_p, v) = a_v \cdot e^{-\frac{(\text{height}(s_p)-b_v)^2}{c_v^2}}
$$

(2)

where for each species $v \in V$, $a_v$ is the value at the highest point of the peak of the curve of the Gaussian distribution, $b_v$ is the average height value, and $c_v$ is an approximation of the standard deviation of the heights for that species.

Fig. 3 (a) Height-map with altitudes ranging between 800 and 1300 meters. (b) Slope map, which varies between 0 and 90 degrees
Finally, the distribution of plants per species \( w(s_p, v) \) has to be normalized to obtain how the height \( w_{\text{height}}(s_p, v) \) conditions a vegetable species \( v \) to be assigned to a particular seed \( s_p \). It is obtained by means of (3):

\[
\begin{align*}
\text{height}(s_p, v) &= \frac{w(s_p, v) - \min(w(v))}{\max(w(v)) - \min(w(v))} \\
\end{align*}
\]

where \( \min(w(v)) \) is the minimum value of \( w(s_p, v) \) for species \( v \) considering all the seeds on the terrain and \( \max(w(v)) \) is the maximum.

Assuming the same reasoning for the slope, the value \( w_{\text{slope}}(s_p, v) \) would be obtained in a similar way to (3), but in this case considering the function \( \text{slope}(s_p) \) that measures the slope of the terrain where the seed is positioned, instead of \( \text{height}(s_p) \) in (2).

5.2 Influence of the grouping of the vegetation species

In a real outdoor environment, it is common to find plant groups of the same vegetable species. In order to simulate these groupings, the method allows for the addition of areas where a specific vegetable species has more likelihood of growing. The number of grouping areas where a vegetable species is more profuse is an input parameter of the method. They are represented by certain points on the terrain, called diffusion points, considered as the center of a circular area where the grouping is allocated. These diffusion points can be positioned randomly in the work presented or can be specified by the user. Figure 4 shows an example where some diffusion points for three different species have been distributed. They have been differentiated by three colors: blue, red, and yellow.

For each seed on the map, the method searches to determine whether it is situated in the area of influence of one or more diffusion points. This value will reach its maximum when it is right in one diffusion point and will be reduced until reaching the limit of its area of influence, with a value of 0 just when leaving it. Equation (4) illustrates the calculation performed to measure the influence of each diffusion point \( \text{diff}_v \) belonging to the species \( v \) on a seed \( s_p \) as follows:

\[
\begin{align*}
\text{diff}_v(s_p) &= \begin{cases} 
1 - \frac{d(s_p, \text{diff}_v)}{r(\text{diff}_v)}, & \text{when } d(s_p, \text{diff}_v) \leq r(\text{diff}_v) \\
0, & \text{otherwise}
\end{cases} \\
\end{align*}
\]

Fig. 4  (a) Diffusion points for three different species, differentiated by three colors (blue, red, and yellow).  (b) Areas of influence associated with the diffusion points
where \( d(s_p, \text{diff}_v) \) is the distance of the seed under consideration \( s_p \) from the diffusion point \( \text{diff}_v \), and \( r(\text{diff}_v) \) is the radius of the area of influence associated to it.

Finally, the influence of each diffusion point set by the user for a vegetation species \( v \) has to be evaluated on a seed \( s_p \) following the expression:

\[
w_{\text{diff}}(s_p, v) = \arg \max_{\text{diff}_v \in V} w_{\text{diff}}(s_p) \tag{5}
\]

### 5.3 Labeling process

The abiotic features and the grouping simulation have to be considered together to perform the labeling process \( w_{sp}(s_p, v) \). Equation (6) shows how this weight is obtained by performing a linear combination between the influence of the terrain, \( w_{\text{height}} \) and \( w_{\text{slope}} \), and the influence of the diffusion points \( w_{\text{diff}} \). The user can modify the degree of importance of each of them by managing several coefficients \( (c_{\text{height}}, c_{\text{slope}}, c_{\text{diff}}) \). By default, the same value is maintained for the terrain features as for the area of influence, i.e., \( c_{\text{height}} + c_{\text{slope}} = c_{\text{diff}} \). If the seed that is being evaluated is outside the area of influence of that species, \( w_{\text{diff}} = 0 \) and the weight is obtained considering only the height and the slope.

\[
w_{sp}(s_p, v) = c_{\text{height}} \cdot w_{\text{height}}(s_p, v) + c_{\text{slope}} \cdot w_{\text{slope}}(s_p, v) + c_{\text{diff}} \cdot w_{\text{diff}}(s_p, v) \tag{6}
\]

The vegetation species \( v \) that will be assigned to that seed \( s_p \) will be the one that gives \( w_{sp}(s_p, v) \) its maximum value.

\[
w_{\text{species}}(s_p) = \arg \max_{v \in V} w_{sp}(s_p, v) \tag{7}
\]

being \( V \) the set of the different vegetation species that are going to be present in the outdoor environment.

### 5.4 Computing neighboring information

Once the initial subset has been labeled, the rest of them have to be processed. In nature, it has been demonstrated that plants can condition vegetable species that grow near them, because of certain facts such as the dissemination of seeds from them. Some studies in the literature have simulated these clusters by means of different techniques. For instance, \([29]\) used kernel deformation to create an algorithm to label species by neighborhood.

In the method presented here, it has been considered that each seed conditions the label of its neighbors to create this influence. Consequently, this has been taken into account in the labeling process: the method uses a \( k \)-nearest neighbor classifier \([13]\), in our case with \( k = 3 \) neighbors.

Once the sub-set has been labeled, the process assesses the rest of the seeds by analyzing the label of the three nearest seeds in the initial set. The label is assigned according to the vegetation species with the majority voting among these three neighboring seeds. This makes it possible to simulate the neighborhood that occurs in a real environment, allowing some plants to condition the species of the nearby seeds.

Figure 5a shows the labels applied to the seeds in the initial set, a different color being assigned to each vegetation species. If the position of a seed is not appropriate for any of the vegetation species to be distributed, it is marked in black and an empty label is assigned to it. Figure 5b illustrates the labels of all the seeds on the terrain after performing the labeling process that takes the neighboring vegetable species into account. This figure shows some areas where a vegetation species has a strong influence, and some others where different vegetation species appear in the middle. This fact simulates the randomness of nature.
Fig. 5  (a) Labeling of the training set of seeds. Three vegetation species (blue, red, and yellow) were considered. Black dots represent seeds with no vegetation species assigned. (b) Distribution of the vegetation species obtained by applying the $k$ nearest neighbor classifier on the rest of the points with $k = 3$

6  Inequality systems to maximize plant representations

The main contribution of this work is the design of inequality systems adapted to each seed in order to compute the size of the plant representation. Once a vegetable species has been assigned to every seed, the size of each plant representation has to be obtained in order to cover most of the terrain. This goal was accomplished by designing a set of inequality systems that consider different (biotic and abiotic) features to estimate the maximum size of each plant without overlapping between neighboring plants. This size is represented by a circle around the seed, similar to the one used in the FON model [4, 5], which determines the maximum area that can be reached by the plant representation. Its radius can be estimated taking into account that the projected plant canopy on the ground is maximized.

This problem can be treated using an optimization technique that establishes relationships between distances with constraints as presented for foams in [25]. In this paper, inequality systems consider some biological data about the vegetable species assigned to the seed in addition to some general constraints:

- The circle of each plant cannot intersect with any other plant. This is not true in nature, where a certain overlapping between the tree canopies does exist, but this overlap will be performed during the final visualization process.
- The radius of the circle of each plant cannot exceed the maximum radius that the vegetation species is biologically capable of reaching.
- If the circle associated with the plant does not reach the minimum radius for the species, that plant cannot survive.

Given a distribution of points that represent the seeds, a distance matrix is constructed. Based on this information, we can estimate the area occupied by a set of plants on a terrain or domain $\Omega$. Let $C_r(s_p)$ be any circle centered on point $s_p$ with radius $r$ and whose area $A$ is bounded by:

$$A(C_r(s_p)) = \pi \cdot r^2$$  (8)
Let \( P = \{s_{p1}, \ldots, s_{pn}\} \) be a set of \( n \) isolated seed positions representing the coordinates where the plants have been planted. Considering \( \Omega \) as the possible locations on the terrain, i.e., \( \Omega \subset \mathbb{R}^2 \), this set \( P \subset \Omega \) is associated with the subset of \( \mathbb{R}^2 \), given by (9).

\[
D_\Omega(P) = \left\{ (r_1, r_2, \ldots, r_n) \subset \mathbb{R}^n \mid C_{r_i}(s_{pi}) \subset \Omega \text{ for } i = 1, \ldots, n \right. \\
\left. \text{and } C_{r_i}(s_{pi}) \cap C_{r_j}(s_{pj}) = \emptyset \text{ for all } i \neq j \right\}
\]

where \( D_\Omega(P) \) corresponds to the solution space to which all the radii found by the method belong. The solution \( D_\Omega(P) \) found from the above equation guarantees that the circular areas of the plants do not intersect each other. These regions called circular areas without intersection (cawi) in set \( P \) are determined according to

\[
\text{cawi}(P, \Omega) = \max_{(r_1, \ldots, r_n) \in D_\Omega(P)} \left\{ \pi \cdot \sum_{i=1}^{n} r_i^2 \right\}
\]

(10)

The following example clarifies the procedure. Let us consider the case of a domain \( \Omega \) where three plants of different vegetation species are defined: \textit{Pinus Sylvestris} (A), \textit{Abies Alba} (B), and \textit{Sorbus Aucuparia} (C). Their maximum radii of growth are 5.6, 5.7, 2.7 meters and their minimum radii are 0.2, 0.6, 0.7 meters, respectively. Thus, we define the distances between plants with values of 7, 8, 9 meters. Figure 6 illustrates these data by means of a graph that relates all these distances between a plant of each vegetation species and the vector \( M \), which represents the maximum radius that each plant can reach.

From the graph, the following system of inequalities can be extracted:

\[
\begin{align*}
   r_A + r_B &\leq 7, \\
   r_A + r_C &\leq 8, \\
   r_B + r_C &\leq 9, \\
   r_A &\leq 5.6, \\
   r_B &\leq 5.7, \\
   r_C &\leq 2.7, \\
   r_A &\geq 0, \\
   r_B &\geq 0, \\
   r_C &\geq 0.
\end{align*}
\]

(11)

Fig. 6 Graph analyzing all the data involved in the example regarding the three plants belonging to the three vegetation species under consideration (A, B, C). Green lines represent the distance between the plants and brown lines represent the maximum radii that each of them can reach until the meeting point \( M \).
This system of inequalities can be identified by $S \cdot r \leq b$, where $S$ is a matrix of zeros and ones that relates the different plants to each other and to the vector $M$ of maximum radii; $r$ is the radii vector to be solved, in this case $(r_A, r_B, r_C)$; and $b$ contains the independent terms of the inequalities.

In this example, the solution that maximizes the $cawi$ value for the three plants is given by (12).

$$cawi(P, \Omega) = (5.6^2 + 1.4^2 + 2.7^2) \cdot \pi \approx 127.58.$$  (12)

On observing the system of inequalities, it can be seen that the constraints do not include the minimum radius that each vegetation species requires in order to live. The reason for this is to favor the growth of large trees rather than small trees, since these provide the $cawi$ value with a greater numerical result. This decision allows plants with a radius lower than their minimum radius to be eliminated in the same way as was performed with a self-thinning model in [29].

Finally, the whole process is repeated a second time so that the space left by the plants that are removed can be covered by the rest of the plants. Figure 7 shows the result of applying this growth method in the terrain used as an example by solving the system of inequalities. As can be observed, the previous points with the same size shown in Fig. 5 now appear re-sized with different magnitudes.

The proposed method is flexible when it comes to generating different plant distributions depending on the user requirements. If required, a more regular distribution could be generated on the terrain, for example by homogenizing the size of the plant population by not allowing the existence of small plants, or simulating the possibility that if a plant reaches its maximum size, it can die of old age and then its space can be occupied by the plants around it.

7 Results

Several experiments were performed using a computer with an Intel Xeon E5 v2 3.70GHz processor. The trees and textures used in the examples belong to the Landscape Automaterial plugin [22]. Unity 3D version 2018.2.7 was used to view the resulting outdoor environments with a realistic visualization.

Fig. 7  (a) The $cawi$ obtained is shown for each seed on the terrain. (b) The detail on the right shows the different sizes of the circles
Three different vegetation species were considered in the experiments: *Pinus Sylvestris*, *Abies Alba*, and *Sorbus Aucuparia* (Fig. 8). They have similar characteristics in terms of the type of ecosystem they need to live in, but conversely they differ in size and in the characteristics of the terrain (height and slope) on which they can grow. The data requirements for each of them were acquired from real data [1, 36, 37] and are summarized in Table 1. This table provides information about the minimum and the maximum radii that a plant of that species can reach, the minimum and the maximum height where it can live, and the maximum slope where that species can appear.

In order to evaluate the influence of the terrain features on the labeling of a seed, the values $a_v$, $b_v$, $c_v$ are required for each species so as to be able to compute (2). These values were obtained from Growth-Grammars [16]. They are shown in Table 2, classified according to the height and the slope of the terrain.

Regarding the diffusion points, the ones that were considered are those shown in Fig. 4. The labeling process was performed following a 3 nearest neighbor classifier, as was previously reported. Finally, the system of inequalities was solved by applying the Simplex optimization technique, since this technique makes it possible to obtain the best solution without constraint in the number of variables, thus allowing these variables to be expanded should it be necessary to do so in the future.

### 7.1 Validation of the method

The method has been tested using a height-map with a resolution of $512 \times 512$ meters. As has been said in previous sections, the relation established between pixels and meters has been $1 \text{ meter} = 1 \text{ pixel}$. Regarding the height of the terrain, a height level difference of 500 meters has been represented between the lowest and highest areas, with a range of values that vary between 800 and 1,300 meters.

Some tests were performed to validate the procedurally created outdoor environments, specific modifications being applied in each execution. Initially, the areas with no vegetation are established by limiting the height values: areas higher than 3,000 meters are considered dry zones. The same map was used in all the experiments to make it easier to notice the visual differences when some parameter values change.

Three different vegetation species were considered and differentiated by colors: *Abies Alba* is represented in red, *Pinus Sylvestris* is shown in blue, and *Sorbus Aucuparia* is shown in yellow. Different parameter variations were tested in order to see how the distribution

![Species of trees used, from left to right: Pinus Sylvestris, Abies Alba, and Sorbus Aucuparia](Fig.8)
Table 1  Data for the three selected vegetation species

| Vegetation          | Radius | Height | Max |
|---------------------|--------|--------|-----|
|                     | min    | max    | min | max | slope |
| **Pinus Sylvestris**| 0.2    | 5.6    | 1,000 | 2,000 | 70   |
| **Abies Alba**      | 0.6    | 5.7    | 700  | 2,000 | 60   |
| **Sorbus Aucuparia**| 0.7    | 2.7    | 0    | 2,300 | 30   |

The minimum and maximum radii of the bounding box that the plant can reach and, regarding the terrain, the minimum and maximum height and the maximum slope where each plant can grow. All these data are measured in meters and the slope in degrees.

changes according to the user’s preferences. All the data involved in the tests are shown in Table 3 and the results obtained can be seen in Fig. 9.

The first test has been called Basic map (Fig. 9a). Initially, the values that weight the parameters involved in the labeling process of the training test were set to \((c_1 + c_2 = c_3)\). They give the same relevance to the terrain characteristics as the areas of influence established by the user. In this case, the values considered were \(c_1 = 0.25\), \(c_2 = 0.25\) and \(c_3 = 0.5\). The visual result shows that the different plants appear grouped by species because the area of influence has played the leading role. As indicated in Table 3, 7,581 seeds were initially distributed, with a spacing between the plants of 5 pixels, but after applying the requirements for growth, 2,145 plants were eliminated. This test required the resolution of a system of inequalities made up of 7,581 inequalities.

The second test, called Height relevance in the table, evaluates the visual results when the influence of the height is increased. The values given to the coefficients in this case were \(c_1 = 0.6\), \(c_2 = 0.2\), \(c_3 = 0.2\). The same distribution and number of seeds were set. These data also produced the same number of inequalities as in the previous test. However, finally, more plants were eliminated: 2,476. Figure 9b illustrates the visual result. Plants of different species appear to be more mixed than in the previous test. This is because all of them have a similar range of height where they can grow, as can be seen in Table 1.

The third test was performed by modifying the range of heights of the terrain. Maintaining the variation of 500 meters, the scale of heights was modified from 1,900 to 2,400 meters. Regarding the data shown in Table 3, it can be observed that on establishing the same number of initial seeds, the number of plants eliminated was quite similar to that of previous tests (2,434). Figure 9c shows the visual results.

On the one hand, an area without any plants can be observed. This area is the highest one on the map and so, by increasing the height range above the maximum height where the three species can grow, an empty area appears.

Table 2  Estimated values used to calculate the number of plants per species depending on the height and the slope ((2))

|        | Pinus Sylvestris | Abies Alba | Sorbus Aucuparia |
|--------|------------------|------------|------------------|
| Height | \(a_v\)          | 44,200     | 6,060            | 2,056           |
|        | \(b_v\)          | 1,041      | 1,553            | 1,396           |
|        | \(c_v\)          | 584.7      | 230.6            | 346.7           |
|        | \(a_v\)          | 46,582     | 2,834            | 1,401           |
|        | \(b_v\)          | 36.60      | 50.21            | 49.01           |
|        | \(c_v\)          | 26.91      | 27.52            | 25.57           |
Fig. 9 Coverage maps obtained by modifying some features. The vegetation species considered are: *Abies Alba* (red); *Pinus Sylvestris* (blue); *Sorbus Aucuparia* (yellow); *Pinus pinea* (magenta); *Quercus robur* (cyan); *Betula pendula* (green)

On the other hand, the plant represented in yellow, *Sorbus*, has a considerably increased number of instances because the maximum height where it can grow is higher than the other two vegetable species. In the low-lying areas, however, the dominant vegetable species becomes *Pinus Sylvestris* (blue), given that its minimum height is lower than that of the red element, *Abies Alba*, and it has a higher slope value than the other two species.

Finally, it can be seen in this figure that despite the fact that there is less area covered by vegetation, the number of eliminated seeds is similar to the rest of the tests. This is
Table 3  Details of the parameters involved in the experiments performed

| Experiment name          | Terrain Resolution | Initial Seeds | Inequality Number | Eliminated Seeds | Spacing (pixels) |
|--------------------------|--------------------|---------------|-------------------|------------------|------------------|
| Basic test               | 512 × 512          | 7,581         | 37,905            | 2,145            | 5                |
| Height relevance         | 512 × 512          | 7,581         | 37,905            | 2,476            | 5                |
| Increasing heights       | 512 × 512          | 7,581         | 37,905            | 2,434            | 5                |
| Increasing spacing       | 512 × 512          | 1,933         | 9,665             | 242              | 10               |
| Doubling no. species     | 512 × 512          | 7,581         | 37,905            | 2,108            | 5                |
| Larger map size          | 1024 × 1024        | 35,969        | 179,845           | 5,428            | 5                |

because the dominant vegetable species has a smaller maximum size than the other two and, therefore, more elements fit in the same area than in the other tests with other dominant species.

Regarding the fourth test, the initial spacing between seeds increased twofold, resulting in 10. This has been called “Increasing spacing”. In this case, as more space is required, a lower number of seeds were distributed: 1,933. This produced only 9,665 inequalities, which, after solving them, meant that 242 plants eventually died. Figure 9d illustrates the visual result of this test. It shows that if the plants have more space available around them, most of them can grow to their maximum size, thus reducing the number of those eliminated.

The fifth test doubled the number of vegetation species involved in the experiment, resulting in 6 vegetation species. The denomination “Doubling no. species” was applied in the table. Figure 9e shows the visual result of the experiment. In this case, the new vegetation species considered are: *Pinus pinea* (represented in magenta), *Quercus robur* (cyan), and *Betula pendula* (green). The characteristics regarding the radius of the crown are obtained from [36] and the range of the terrain features (height and slope) were considered to be similar to the three vegetation species shown in Table 1. Figure 9e shows how the competition increases and the size of the groups of vegetation species are smaller.

Finally, the last test was performed, doubling the resolution of the terrain. In this case, a resolution of 1024 × 1024 meters was established. On considering a double-sized terrain and maintaining the same spacing between seeds, 5 meters, the number of seeds distributed increased to 35,969, and so the number of inequalities reached 179,845. In the same way, the number of plants that were finally deleted increased to 5,428. Figure 9f shows the visual results obtained by applying the method on this large map. This image shows that there is a greater number of elements, but the competition between them is not very different from the results obtained using the initial map size.

### 7.2 Computational time analysis

The computation time of the method is conditioned by the number of inequalities that have to be solved in order to perform the space-filling process. As stated in Section 6, they are stored as $S \cdot r \leq b$, where $S$ is the matrix storing 1 or 0, $r$ is the radii vector to be solved, and $b$ represents the independent terms of the inequalities. Some of the inequalities considered control the space each plant occupies according to its maximum radius, while others control the distance from that plant to the rest of the plants distributed over the terrain. On analyzing
the size of this matrix $S$, it is observed that given a set of $n$ plants, the number of inequalities that should be incorporated if all possible distances are considered is determined by

$$2 \cdot n + \frac{n \cdot (n - 1)}{2}$$

(13)

However, the number of inequalities can be reduced by considering that the resulting value of the plant radius is directly related to the distance to its neighboring plants instead of to all of them. This approach is a critical aspect in order to reduce the number of inequalities and make the method effective at the computational level.

Several tests were carried out to choose the appropriate number of neighbors without producing overlapping of crowns between plants. According to their results, for each vegetation species only the distances related to 5-nearest neighboring plants are required. This approximation reduces the number of inequalities to the value shown in (14). Nevertheless, it has been demonstrated experimentally that this number can be reduced by up to 40% if the distances that are repeated between neighbors are eliminated.

$$2 \cdot n + 5 \cdot n$$

(14)

Figure 10 shows an example of the distribution of plants of the vegetation species *Pinus Sylvestris*. The terrain was considered to be a $40 \times 60$ meter rectangle where the seed placement followed a uniform distribution. Figure 10a illustrates the results of this distribution, where all distances are considered, while in Fig. 10b only the 5-nearest neighbors are taken into account. The Fig. 10b shows that with only the 5-nearest plants, the method associates to each seed a circular area that does not intersect with any other and produces a good growth proposal.

Finally, in order to analyze the computational cost regarding the number of plants, some populations of *Pinus Sylvestris* were placed over an area of $512 \times 512$ meters using this assumption of considering the 5-nearest neighbors. Figure 11 analyzes the results of this experiment. It can be observed how the time in seconds tends to grow exponentially with the number of seeds distributed over the terrain.

### 7.3 Visual results

The process of obtaining the final terrain is illustrated in Fig. 12, in accordance with the experiment called *Basic test*. The terrain has been textured to show the placement of the different vegetation species and the difference in growth between the plants. Colored circles have been used to offer a visual representation of the relationship between the coverage
Fig. 11 Evolution of computing costs with respect to the number of trees

Fig. 12 Visual process to obtain the final terrain with trees from the height-map

(a) Top view of the result.  (b) Perspective view of plant covering.  (c) Final result with trees.
obtained for the map under consideration and the final terrain with trees. Figure 12a displays a top view of this coverage map and Fig. 12b shows a perspective view, which allows the characteristics of the terrain to be observed visually.

Lastly, Fig. 12c illustrates the final visual result, by changing the circles for the corresponding trees. The realistic model where each plant is represented allows us to value how the characteristics of the terrain have conditioned the final results. Apart from the initially established areas without vegetation, the feature that has determined the lack of vegetation was the slope. Height was not relevant because the maximum height considered in the experiments (1,300 meters) has allowed growth of all the species in the test.

In order to verify the good performing of the method with real terrain, a similar simulation has been carried out using three plant species and a height map obtained from a GIS (see Fig. 13).

Figure 14 illustrates the results in more detail. The Landscape Auto-material plugin has been used in the representation, so the ecosystem generated can be observed better. This Figure shows three captures of the visual results produced at different distances. Figure 14a shows the different vegetation species in the outdoor environment at a long distance. Moving closer to the camera, Fig. 14b shows the scenario at a medium distance. Here, the lack of vegetation can be seen in areas with high slopes. Finally, Fig. 14c illustrates a shot from the ground, where the good quality of the visual results can be seen.

(a) Top view of the result.  (b) Perspective view of plant covering.  (c) Final result with trees.

Fig. 13  Tree distribution on a map of the Benicassim area (Castellón, Spain)
8 Conclusions

This article presents a procedural modeling method for natural environments that obtain the size of each plant representation using inequality systems. They consider biological restrictions, so if the size obtained does not reach the minimum to live, this plant is eliminated. Following this process, the proposed method obtains realistic visual results. The inequalities that control the size of each plant can be customized, adapting the final representation to the requirements of the final application. For instance, overlapping has been avoided in the work presented, although this restriction can be changed to allow canopies to mix, as occurs in nature.

The ecosystem modeling method proposed here starts by performing an automatic positioning of seeds, which are later labeled as a particular vegetation species. Different variables condition these labels, mainly based on the characteristic of the position of the...
seeds on the ground and on the biological requirements of each vegetation species. The labels of the $k$-neighboring seeds are also considered to represent the typical grouping of plants present in nature. Users are allowed to partially condition the final labeling process by establishing certain areas where these groups of plants have to be allocated.

The final realistic representation is obtained by performing the calculation of the final plant representation that grows from the seeds solving the inequality systems. Maximum coverage of the terrain is achieved by allowing the plants to reach the maximum size, and eliminating them if they do not reach the minimum size. Intersection between the volumes assigned to each plant is not allowed in this work, these restrictions being included as inequalities, although some branches can reach other plants in the final realistic representation.

Results show that the outdoor environment generated meets the requirements given for each vegetation species. Furthermore, the visual results are realistic and natural due to the heterogeneity in the size and position of the plants, the grouping achieved between species, and the fact that some elements of another species grow among them.

As future work, this methodology could be extended with the aim of covering large areas of vegetation. The use of Wang-tiles with KD-trees [43] has been considered as a way to solve this problem. These multilevel data structures partition the space into instantiated representations where each tile is filled with PDD. It could easily be extended to use Wang-tiles in order to generate patterns of plant distribution.

Another line of research to work on is to improve the method by adding other biological characteristics of the species, such as soil moisture, temperature, and the position of the sun. In addition, the possibility of adapting the plant representation according to the age that is appropriate to each size is being analyzed. Only one representation has been used in this work, but more biological precision would be obtained if different plant models representing different levels of growth were used for each vegetation species.

**Acknowledgments** This work was supported by the Spanish Ministry of Science and Technology (Projects PID2019-106426RB-C32/AEI/10.13039/50110011033 and RTI2018-098651-B-C54/AEI/10.13039/50110011033) and the Universitat Jaume I research projects (UJI-B2018-56) and (UJI-B2018-44).

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

**References**

1. ARBOLAPP ARBOLAPPES (2021) Available online: http://www.arbolapp.es/. Last accessed July 2021
2. Aikio S (2004) Competitive asymmetry, foraging area size and coexistence of annuals. Oikos 104:51–58
3. Aikio S, Pakkasmaa S (2003) Relatedness and competitive asymmetry -Implications for growth and population dynamics. Oikos 100:283–290. https://doi.org/10.1034/j.1600-0706.2003.11814.x
4. Alsweis M, Deussen O (2005) Modeling and Visualization of Symmetric and Asymmetric Plant Competition. In: Proceedings of the First Eurographics Conference on Natural Phenomena, pp 83–88
32. Mi LM, Hädrich T, Scheeffczyk J, Michaels DL, Pirk S, Lubicki WP (2019) Synthetic Silviculture: Multi-Scale modeling of plant ecosystems. ACM Trans Graph 38.4
33. Onrust B, Bidarra R, Rooseboom R, van de Koppel J (2015) Procedural Generation and Interactive Web Visualization of Natural Environments. In: Proceedings of the 20th International Conference on 3D Web Technology. Web3D ’15. ACM, Heraklion, pp 133–141. isbn: 978-1-4503-3647-5. https://doi.org/10.1145/2775292.2775306
34. Onrust B, Bidarra R, Rooseboom R, van de Koppel J (2017) Ecologically Sound Procedural Generation of Natural Environments. In: International Journal of Computer Games Technology, pp. 1–17. https://doi.org/10.1155/2017/7057141
35. Perlin K (1985) An Image Synthesizer. SIGGRAPH Comput Graph 19.3:287–296
36. Pretzsch H, Biber P, Uhl E, Dahlhausen J, Rötzer T, Caldentey J, Koike T, van Con T, Chavanne A, Seifert T, Toit B, Farnen C, Pauleit S (2015) Crown size and growing space requirement of common tree species in urban centres, parks, and forests. Urban Forestry Urban Green 14.3:466–479. https://doi.org/10.1016/j.ufug.2015.04.006
37. Pretzsch H, Matthew C, Dieler J (2012) Allometry of tree crown structure relevance for space occupation at the individual plant level and for Self-Thinning at the stand level. In: Growth and defence in plants: Resource allocation at multiple scales. Springer, pp 287–310
38. Prusinkiewicz P, Lindenmayer A (1996) The Algorithmic Beauty of Plants. Springer, Berlin. isbn: 0387946764
39. Rebollo C, Remolar I, Chover M, Ramos JF (2004) A comparison of multiresolution modelling in Real-Time terrain visualisation. In: Computational science and its applications -ICCSA 2004. Springer, Berlin, pp 703–712
40. Rietkerk M, Van de Koppel J (2008) Regular pattern formation in real ecosystems. Trends Ecol Evol 23.3:169–175. issn: 0169-5347. https://doi.org/10.1016/j.tree.2007.10.013
41. Smelik RM, Tutenel T, Bidarra R, Bedrich B (2014) A Survey on Procedural Modelling for Virtual Worlds. Comput Graph Forum 33.6:31–50
42. Temmerman S, Bouma TJ, Bidarra R, Van de Koppel J, Van der Wal D, De Vries MB, Herman PMJ (2007) Vegetation causes channel erosion in a tidal landscape. Geology 35.7:631–634. https://doi.org/10.1130/G23502A.1
43. Weier M, Hinkenjann A, Demme G, Slusallek P (2013) Generating and Rendering Large Scale Tiled Plant Populations. J Virt Reality Broadcast 10.1:1860–2037
44. Zegers CD (1997) Ecología forestal: el bosque y su medio ambiente. Colección Nueva técnica: Universidad Austral de Chile. Editorial Universitaria

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.