

# A Structural Growth Model of the Invasive Weed Species Yellow Starthistle, *Centaurea solstitialis* L.

<sup>1</sup>Thornby, D., <sup>2</sup>J. Garren, <sup>2</sup>R. Carruthers, <sup>2</sup>D. Spencer and <sup>2</sup>J.P. Weber

<sup>1</sup>Queensland Department of Primary Industries and Fisheries <sup>2</sup>USDA Agricultural Research Service;  
Email: david.thornby@dpi.qld.gov.au

*Keywords: Centaurea solstitialis, yellow starthistle, structural model, LStudio, L-systems.*

## EXTENDED ABSTRACT

Yellow starthistle (*Centaurea solstitialis* L.) is a summer-flowering annual plant of the family Asteraceae, and a significant weed of roadsides, pasture and grasslands in California. It is also distributed, though with less economic and ecological significance, throughout the US and in other temperate parts of the world. Yellow starthistle grows vigorously during the dormant period for most native species in its grasslands habitat, is able to access deep soil moisture reserves, and being particularly spiny, reduces land value for both animal fodder and human recreation. It is a prolific seed producer and seed spreads readily through contact with animals, people, and vehicles.

As with many weeds, yellow starthistle's structure and growth patterns have been relatively little-studied. It is hoped that studying the ontogenesis and morphogenesis of yellow starthistle will help uncover information that can assist in improving control strategies, as well as providing a visualisation tool for disseminating detailed knowledge about this plant's growth and development to stakeholders in a readily understood way.

To produce a visual simulation of the morphogenesis of yellow starthistle, a modular, component-based approach was taken, using L-systems modelling. First, a base model of the plant's bolting, branching and flowering behaviour was constructed using casual observations of real plants. Second, the model was parameterised to fit data on morphology of yellow starthistle plants taken from three different locations in northern California: trial sites at Cache Creek, Bear Creek, and Putah Creek. Plants were collected throughout the growing season, dried, and later digitised using a sonic digitiser. Data from the digitiser were analysed and relationships between time and various aspects of morphology were devised. The base model was adjusted in light of these data, and parameterised to fit mean values for the different field sites from which data was collected.



**Figure 1.** Virtual yellow starthistle

Stochasticity was added to the model, so that it reproduces a range of plants of different sizes and morphologies, within the range of likely sizes observed in the data for each field site.

While the initial model explored possible mechanisms behind yellow starthistle's growth characteristics, the finished model is an empirical (but stochastic) representation of yellow starthistle's growth in three different field sites. The model therefore demonstrates the morphogenesis of yellow starthistle in three different sets of environmental conditions, in a readily understood way - as a three-dimensional, dynamic simulation. The model has potential for use as a morphogenetic template to which simulations of responses to stimulus (such as damage to reproductive structures) and interaction with potential biocontrol agents could be added, as well as for exploring in greater detail the morphogenetic effects of particular environmental variables in isolation.

## 1. INTRODUCTION

While there are an increasing number of models investigating and depicting structural growth for several significant crop plants, such as cotton (Hanan and Hearn 2003), maize (Fournier and Andrieu 1999), and rice (Watanabe *et al.* 2005), morphogenetic models of non-crop plants, and in particular of weeds, are less common. This is despite the significant economic and ecological importance of weeds around the world.

### 1.1. Yellow Starthistle

Yellow starthistle (*Centaurea solstitialis* L.) is a non-indigenous weed in the United States, originating in a wide area in Eurasia (Maddox 1981). It has successfully invaded large areas of grazing, horticultural, amenity and roadside land in California, as well as large areas of native grassland. Over 6 million ha of land in California has been invaded by yellow starthistle (Pitcairn *et al.* 1998) and it is well-established throughout the continental US and in temperate areas around the world. It is a summer-flowering annual of the family Asteraceae, the capitula having yellow florets and sharp spines attached. After germination, it produces a rosette before progressing into a bolting phase, wherein it sends up an often highly branched structure with much smaller leaves and a variable number of capitula. Spread of the plant from one area to another is primarily due to human and animal activity, as the seeds are not readily spread by wind (Roche 1992).

Yellow starthistle has proven successful in northern California, as it produces a large number of seeds (often 80 or more per capitulum; that is, 100,000 or more on a large plant and around 50 million seeds per heavily-infested ha; di Tomaso *et al.* 1998), grows rapidly during the summer dormancy period of native grassland plants growing in the same area, and has a high ability to access stored soil moisture (di Tomaso *et al.* 2000).

Potentially useful control methods for yellow starthistle include mechanical and cultural controls (mowing, grazing, burning, hand removal), biocontrol, particularly using insects that attack flowers and seed heads, and a variety of chemical options. In order to manage yellow starthistle populations successfully, an integrated selection of these methods is needed, along with more information on aspects of the plant's growth in response to both control methods and environmental cues and conditions. Here, we have used a structural modelling approach to simulate the plant's growth in a number of different field sites, with the aim of investigating how different

environmental conditions (represented in this study by using several different field locations) affect the plant's morphogenesis.

### 1.2. Structural modelling with L-systems

The L-system formalism (Lindenmayer, 1968; Prusinkiewicz and Lindenmayer, 1990) for modelling the growth of modular organisms (including chains of bacteria and plants) was devised initially by Aristid Lindenmayer. The formalism is inherently suitable for modelling plant development, because it simulates modular organism components (such as leaves, internodes, meristems, flowers and fruits) as a set of symbols arranged in a branching topological string, in much the same way that plants, being modular, (Room *et al.* 1994) consist of a set of modules or organs arranged in relatively predictable topologies. L-systems are a rule-based algorithmic system of modelling, where the addition of new components is done in parallel. Several software packages are available for programming L-systems. LStudio (University of Calgary) provides tools for constructing L-system models using the programming language cpfg (plant and fractal generator with continuous parameters) (Hanan, 1997).

Data collection and model development were undertaken to some extent concurrently, although the approach taken in constructing the model was to produce qualitatively approximate morphologies using visual observations of real yellow starthistle plants first, and then to parameterise the qualitative model using the collected data. Visual representations of stem cross-section, leaf shapes, and bud and flower shapes, were produced with reference to plant samples and photographs. Where the quantities and positioning of branches and reproductive structures were concerned, the data (as plant architecture data files) were useful in reconstructing this species' qualitative strategies of branching, flower production and placement, which was rather too complex over time to be apparent from casual observations.

## 2. DATA COLLECTION AND ANALYSIS

Data for construction and parameterisation of the model was taken from a pool of data collected using a sonic digitiser (Polhemus Fastrak) and Floradig software (University of Queensland) (Hanan and Room, 1997) from a large number of dried plants. The plants used in the dataset for constructing this model were collected in the field during 2001 in three different locations in northern California: field trial sites at Cache Creek, Putah Creek, and Bear Creek. This is part of a larger dataset that includes other plants from these and other sites. Twenty-five plants in the bolting phase of growth were collected at approximately seven-day intervals over the growing season in 2001

(between 29 May and 23 August) and were dried for later digitising. The digitiser was used in conjunction with Floradig software to collect and organise information on the spatial and topological position of leaves, internodes, and reproductive structures. Using the raw data points collected during digitising, Floradig calculated lengths of internodes and leaves, the total height of the plant, and angles of attachment of branches and leaves.

The data were summarised using SAS, Microsoft Access, and Excel. In order to represent yellow starthistle's pattern of growth, the key data used (for each field site) were: mean numbers of each relevant plant organ present over time; mean numbers of branches and metamers per branch at each level of the topological hierarchy present over time; and mean fully-expanded sizes of internodes and leaves at each level of the topological hierarchy.

Curves were fitted to the data for numbers of branches at each branching level over time, in order to describe how the branching structure develops numerically. These curves were then reproduced with LStudio's cubic spline function editor (see below).

### 3. L-SYSTEM PROGRAMMING

The structural model of yellow starthistle consists of a set of L-system symbols (Table 1), a set of rules (called productions) that modify the structure iteratively, and a set of drawing instructions that produce a dynamic visual simulation of the growth of a single plant (the homomorphism).

Thus, a string of symbols representing a short part of a bolting-phase yellow starthistle stem might be:

Ib Lb [B] Ib Lb [Ib Lb[B]] Ab

The square brackets indicate the start and end of branches in the string.

Some of the symbols in the L-system have one or more parameters associated with them. The parameters are used to keep track of various pieces of qualitative information about particular components, including age in days, a number index indicating a component's position of attachment on a shoot, and the level of branching occupied by a component (where the original main stem is level 1) (Table 1). Symbols only keep track of those parameters necessary for determining the effect of the model's growth rules.

The model proceeds on a daily time step. Accumulated days since the beginning of the growing season are used in predicting the growth of the plant (that is, number, size, and placement of internodes, leaves and reproductive structures) according to relationships derived from the collected data. The starting state for the model is a single rosette-stage apical meristem. In order to simulate growth dynamically, L-systems apply a set of rules (productions) to the current string of symbols, and the resulting string replaces the current one at the end of the step. The productions used in this model are summarised in the statements given below, although for clarity the programming syntax used to implement the model is not reproduced in its original form.

#### 3.1. Apical growth

The plant's main stem apex proceeds through two stages: the rosette stage and the bolting stage. The rosette stage apex rules can be summarised as follows:

- If the current number of rosette metamers is less than half the predefined total number of rosette metamers, produce a new entire rosette leaf, compressed rosette internode, and rosette apical meristem.

**Table 1.** Symbols used in the yellow starthistle L-system, plant components they represent, and parameters associated with each symbol

Symbol	Component represented	Associated parameters
Lr	First (entire margin) rosette-stage leaf	Age
Ls	Second (serrated margin) rosette-stage leaf	Age
Lb	Bolting-stage leaf	Age
Ar	Rosette-stage apical meristem	Position in stem
Ab	Bolting-stage apical meristem	Branching level, position in stem, total # of metamers
B	Axillary meristem	Branching level
Rb	Reproductive meristem	Age
Rr	Reproductive structure (capitulum)	Age
Ir	Rosette-stage internode	
Ib	Bolting-stage internode	Age, branching level

- If the current number is more than half but below the total number of rosette metamers, produce a new serrated rosette leaf, compressed internode, and rosette apical meristem.
- If the current number is equal to the total number of rosette metamers minus one, produce a new serrated leaf, compressed internode, and a bolting stage apical meristem.

The bolting stage apical meristem rules (which apply to the main stem and branch apices) can be summarised as follows:

- If the current number of metamers on this axis is less than the maximum number for this axis (Table 2), produce a new internode, leaf, axillary meristem and apical meristem. Pass information on branching level and maximum number of metamers for this branch to the new axillary and apical meristems.

**Table 2.** Number of metamers produced per branch at each level of branching simulated

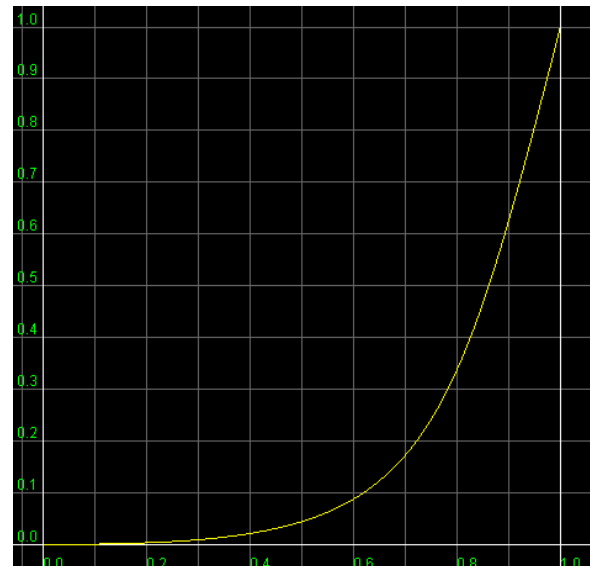
Axis Level	Cache Ck	Bear Ck	Putah Ck
2	8	7	9
3	7	7	6
4	7	6	6
5	8	3	
6	4		

These are relatively simple rules, but are able to reproduce realistic branch lengths. The numbers of metamers produced per branch are whole numbers rounded up from mean numbers of metamers per branch found in the data. The model can be set to run with each branch taking the exact value for axis level and location, or with a random number of branches being the above value plus or minus a user-defined random factor.

### 3.2. Branching behaviours

It is the number and position of branches, levels of branching, and their rate of appearance over the growing season, that are most important for reproducing yellow starthistle's convoluted morphological pattern. In order to reproduce realistic branching patterns, the model makes extensive use of LStudio's cubic spline functions editor (Figure 2). For each branching level, and each field site, a separate function was created describing the relationship between time of year and mean number of branches present. Branches may be initiated in the model up to the 6<sup>th</sup> level of

branching for the Cache Creek site (where the main stem or primary axis is level 1, branches or secondary axes are level 2, sub-branches or tertiary axes are level 3, and so on), the 5<sup>th</sup> level of branching for the Putah Creek site, and the 4<sup>th</sup> level of branching for the Bear Creek site. Around one quarter to one third of the real plants produced branches at the top level for each site. Very few produced greater degrees of branching than this for each site, though one plant was collected from the Cache Creek site with branching up to nine levels.



**Figure 2.** LStudio's cubic spline editor, showing curve for Cache Creek branches per plant over time, at branching level 4. The x-axis represents time proportional to the length of the growing season (where 1.0 represents the end of the growing season), and the y-axis represents the proportional number of branches (where 1.0 represents the full complement of branches for a given axis level)

Each curve was designed to reproduce the curves found in the data for each relevant characteristic, as described above. Each curve is drawn between 0.0 and 1.0, and when multiplied by (that is, scaled to) mean values for maximum number of branches per plant at a particular axis level for a particular field site (Table 3), the function gives a current predicted number of branches at that particular combination

**Table 3:** Mean maximum number of branches produced per plant at each level of branching simulated

Axis Level	Cache Ck	Bear Ck	Putah Ck
2	6.2	5.0	5.0
3	14.0	9.5	6.0
4	19.7	9.5	2.0
5	11.0	4.8	
6	4.3		

of date, axis level and field site. This value is used in determining whether a given axillary bud develops into a branch on a given day. The scaling value can also be determined stochastically, with some (user-defined) variation around the means in Table 3.

The set of productions that determine the behaviour of the axillary meristems (and therefore the branching behaviour of the plant) can be summarised as follows:

- Determine the proportion of the total mean number of branches that are predicted to be present at the current date in the model, by referring to the appropriate curve for the current field site being simulated.
- If this proportion is greater than the current actual number of branches, generate a random number between 0 and 100 and compare it to the probability of a particular meristem initialising a new branch in a particular step.
- If the random number is lower, determine the number of metamers to be produced in the new branch (by referring to a lookup array for the current field site being simulated), and produce a new branch apical meristem.

The fact that each axillary meristem is tested against this rule in each step means that the plant tends to produce more basal branches, and that the basal branches tend to appear earlier than more distal branches. This is broadly representative of real yellow starthistle plants, and the numerical propensity for this characteristic can be adjusted to some degree by changing the branch initiation probability figure used in the rule.

### 3.3. Leaves and internodes

Leaves are represented with a single bicubic surface (Prusinkiewicz and Lindenmayer, 1990) (for entire rosette leaves and bolting-stage leaves) or a string of smaller bicubic surfaces (for serrated rosette leaves). The internodes of yellow starthistle stems are winged (alate); this appearance was reproduced by modelling the stems as generalised cylinders (Prusinkiewicz *et al.*, 2001) with an appropriately-shaped closed contour used for the cross-section of the cylinders.

A simple version of leaf and internode expansion is implemented in the model, as data for expansion of individual internodes and leaves was not collected as part of the dataset used in developing the model. The rules for both internodes and leaves of all types work as follows:

- Determine the proportion of the current age of the leaf or internode to the age at which the component would be fully expanded.
- Multiply the maximum size of the component by the current age proportion, and set this value as the current size of the component.

### 3.4. Reproductive structures

Yellow starthistle, being of the family Asteraceae, produces a typical daisy capitulum, with a disk of apetalous yellow florets subtended by a structure of many bracts. In the case of yellow starthistle, the bracts also produce a set of sharp, tough spines.

These morphological characteristics are simplified in the model to a set of bicubic surfaces for the bracts, arranged around the base of a yellow cone for the florets, with a row of generalised cylinders forming the spines.

In yellow starthistle, the reproductive structures are held at the end of each branch: that is, branch growth terminates with the production of a flower bud, which develops into a capitulum. This morphogenetic pattern is clearly apparent in the digitiser data for each plant. The mechanism by which an apical meristem 'decides' to switch to reproductive function is not clear. In the model, branch developmental length is predetermined and parameterised from the data, as shown in Table 2. An addition to the rules for apical meristem development provides for flower development, making the whole bolting apex rule thus:

- If the current number of metamers on this axis (Table 2) is less than the maximum number, produce a new internode, leaf, axillary meristem and apical meristem. Pass information on branching level and maximum number of metamers for this branch to the new axillary and apical meristems.
- If the current number of metamers on this axis is equal to the maximum number, produce a flower bud and cease further apical growth.

Assuming that the number of branches to be produced by the model for any particular field site is an accurate reflection of the real plant data, numbers of capitula produced by this method should likewise be accurate.

Capitulum age is of particular significance for our study of yellow starthistle, as the activity of various proposed and current biocontrol agents depends on the state of individual capitula and the size of the population of capitula that are in suitable stages of growth at any given time.

The state of any given capitulum, broadly classifiable as buds, flowers, and seed heads in the

current model, but more finely classified into different stages of each of these states in the data, is determined in the model by its age. From this, populations of capitula in each state at any time can be returned by the model.

#### 4. RESULTS

The model produces a range of plants of somewhat different morphologies for each field site (Figure 3 a-c). The range of plants produced represents the ‘average’ yellow star thistle plant with a good level of accuracy. The visual output of the model can be rotated freely, as it is a three-dimensional object. Any view can be saved as a picture file, and the model can also output data on numbers of internodes, branches, buds, flowers, and seed heads (Table 4). Various calculations on these and other values could also be performed within the model if desired, and the output saved for display or analysis.

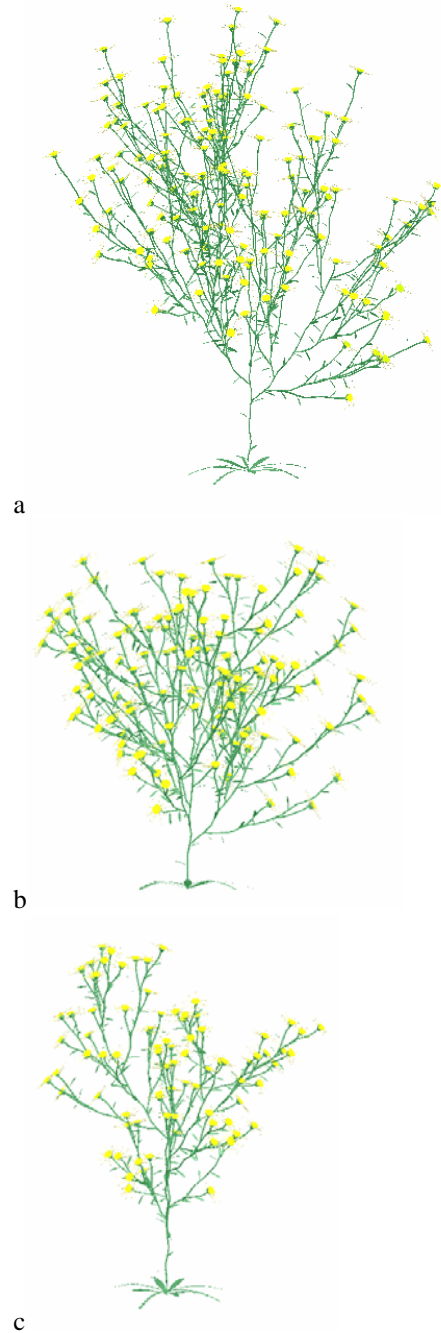
**Table 4.** Sample model output: number of internodes and branches produced by the model over time for three sites: CC- Cache Creek; BC- Bear Creek; PC – Putah Creek.

Day	No. of internodes			No. of branches		
	CC	BC	PC	CC	BC	PC
10	9	9	9	0	0	0
30	58	47	46	7	6	4
60	132	100	92	15	12	9
90	287	185	121	47	30	13

#### 5. CONCLUSIONS

The L-system model of yellow starthistle provides a three-dimensional simulation of the weed’s growth in three different sets of growing conditions. The simulated plants reproduce realistic ‘average’ plants in terms of numbers of branches, internodes, and reproductive structures. However, the range of plant sizes does not fully reproduce the range of real plant morphologies represented in the data. This is particularly the case in terms of branching. The number of branches produced per plant represents the average plant, but the means in the data are in several cases made up of individuals with a large range of branch numbers. For example, for the Putah Creek plants collected on 1 August 2001, at branching level 4, 13 plants had no branches, while one plant had 86 branches; the mean value for the 25 plants collected there on that date was 6.96 branches at that level. The model does not currently produce plants at the extreme upper range of real values found in the data. The more significant consideration is whether or not it is appropriate to simulate the whole of a very divergent range of individuals, or to focus on producing simulations that are reliably representative of ‘average’ individuals. Very large

individuals in the dataset probably grew in areas with markedly less competition and/or more resources than the average for a given field site.



**Figure 3.** Virtual yellow starthistle plants: (a) Cache Creek data; (b) Bear Creek data; (c) Putah Creek data

The meristem-activation method used in the model does not simulate the plastochron (the average time between the appearance of successive metamers) of individual axes; rather, the branch-activation algorithm approximates the whole-plant mean plastochron. The assumption that this is a

reasonable method of controlling the development rate would require more detailed data (most likely, data collected from living individuals at daily intervals) to test and, if necessary, alter.

As it stands, the model is a useful tool for investigating and demonstrating, in a visual format, the dynamic growth of this weed. It captures some of the variation in plant sizes within and between particular field sites. It also forms a template for future modelling and morphological work in this species.

Avenues for improvement and extension of the model could be in areas such as adding parameter values for a wider range of environmental conditions, represented by a range of more widely variable locations, perhaps with the aim of adding greater (mechanistic) detail in how the plant responds to changes in individual environmental variables such as nutrition, soil moisture and temperature. Biomass calculations could be performed by the model given sufficient data for correlating organ sizes with dry weight. Adding greater detail in the treatment of classifications of reproductive structures may also assist scientists working in biocontrol of the weed, in that this would improve the model's ability to provide dynamic information on flower and seed head populations in categories of interest, over the life of the plant. A field of many stochastically different virtual yellow starthistle plants could add greater depth to the model's usefulness for scientists involved in biocontrol research in this weed. Adding virtual insect populations would further improve such a model.

## 6. ACKNOWLEDGEMENTS

This work was carried out as part of a postdoctoral project at the University of California, Davis. The project was sponsored by the US Department of Agriculture; digitising and plant collection were performed by Justin Weber, Carol McGugin, Marie Franc, and Lauren Bingham at the USDA-ARS labs at Albany, CA.

## 7. REFERENCES

- Di Tomaso, J.M., Benefield, C. and Kyser, G.B. (1998), Reproductive biology of yellow starthistle (*Centaurea solstitialis*) maximizing late season control. In: Kelly, M., Wagner, E. and Warner, P. eds. *Proceedings, California exotic pest plant council symposium* 4:58.
- Di Tomaso, J.M., Kyser, G.B., Orloff, S.B. and Enloe, S.F. (2000), Integrated approaches and control option considerations when developing a management strategy for yellow starthistle. *California agriculture* 54 (6):30-36.

- Fournier, C. and Andrieu, B. (1999), ADEL-maize: an L-system based model for the integration of growth processes from the organ to the canopy. Application to regulation of morphogenesis by light availability, *Agronomie* 19:313-327.
- Hanan, J. (1997), Virtual plants – integrating architectural and physiological models, *Environmental modelling and software*, 12(1): 35-42.
- Hanan, J.S. and Room P.M. (1997), Practical aspects of virtual plant research, in Michalewicz, M. T. (ed) *Plants to ecosystems. Advances in computational life sciences*, CSIRO, Melbourne, pp 28-44.
- Lindenmayer, A. (1968), Mathematical models for cellular interactions in development, Parts I and II. *Journal of theoretical biology* 18, 280–315.
- Maddox, D.M. (1981), Introduction, phenology, and density of yellow starthistle in coastal, intercoastal, and central valley situations in California. *ARR-W-20*, pp. 1-33. USDA-ARS.
- Pitcairn, M.J., O'Connell, R.A. and Gendron, J.M. (1998), Yellow starthistle: survey of statewide distribution. Pages 64-66, in: D.M. Woods, ed. *Biological control program annual summary, 1997*. California Department of Food and Agriculture, Plant Health and Pest Prevention Services, Sacramento, CA.
- Prusinkiewicz, P. and Lindenmayer, A. (1990), with Hanan, J.S. Fracchia, F.D., Fowler, D.R., de Boer, M.J.M. and Mercer, L. *The algorithmic beauty of plants*. Springer-Verlag, New York.
- Prusinkiewicz, P., Muendermann, L., Karwowski, R., and Lane, B. (2001) The use of positional information in the modeling of plants. *Proceedings of SIGGRAPH 2001* (Los Angeles, California, August 12-17, 2001), pp. 289-300.
- Roché, B.F., Jr. (1992), Achene dispersal in yellow starthistle (*Centaurea solstitialis* L.). *Northwest science* 66:62-65.
- Room, P.M., Maillette, L. and Hanan, J.S. (1994), Module and metamer dynamics and virtual plants. *Advances in ecological research* 25: 105-157.
- Watanabe, T., Hanan, J.S., Room, P.M., Hasegawa, T., Nakagawa, H. and Takahashi, W. (2005), Rice morphogenesis and plant architecture: measurement, specification and the reconstruction of structural development by 3D architectural modelling. *Annals of botany*, 95: 1131-1143.