

Spatially-explicit competition indices and the analysis of mixed-species plantings with the Simile modelling environment

Jerome K Vanclay

Southern Cross University, Lismore NSW 2480, Australia

Tel +61 2 6620 3147, Fax +61 2 6621 2669, JVanclay@scu.edu.au

Abstract

Individual-based modelling and analysis of forest experiments has been made more accessible to researchers with the advent of modelling environments like Simile from Simulistics.com. Individual-based analyses of tree growth data offer insights not possible with plot-based analyses, especially when the original experimental design has been compromised by mortality or other unforeseen events. The paper illustrates how Simile can be used for individual-based analyses of mixed plantings, and how it can be used to explore the consequences of the resulting statistical models. A mixed-species planting of *Eucalyptus pellita* and *Acacia peregrina* is used to illustrate possibilities.

Key words: Simile, Vensim, Powersim, Stella, systems dynamics, individual-based modelling, distance-dependent competition index, mixed-species, *Eucalyptus*, *Acacia*

Introduction

Plot-based analyses of replacement series trials remain the norm for mixed-species research, despite known limitations and demonstrations of better alternatives (Connolly *et al.*, 2001; Park *et al.*, 2003; Vanclay, 2006). Individual-based analyses of tree growth data offer insights not possible with plot-based analyses, especially when the trial has suffered mortality or other disruptions to the original design (Bristow *et al.*, 2006). In the past, the requirement for tedious data preparation and specialist computer skills (Huston *et al.*, 1988; Grimm and Railsback, 2005) have hampered

widespread use of individual-based modelling (Grimm, 1999; McIntosh *et al.*, 2005), but modern packages such as Simile (Muetzelfeldt and Massheder, 2003) make the approach more accessible. Many researchers remain unaware of the utility of such packages in facilitating analyses of competition at the level of individual trees. This paper examines three propositions, that

1. most mixed-species analyses continue to use simple indices such relative yield or analyses of variance, rather than competition indices and response surface analyses;
2. packages like Simile make the calculation of competition indices more accessible and efficient; and that
3. response surfaces fitted data from mixed-species trials may be useful to refine operational prescriptions, additional to their role in analysing experiment data.

The paper also demonstrates the ease of conducting detailed individual-based analysis of inter-specific competition data.

Use of Competition Indices in Mixed-species trials



Evidence for the first proposition is easy to collect, but hard to interpret. Reviews (such as Forrester *et al.*, 2006) cite many publications that report relative yield, but few that report competition indices or response surfaces. This may be attributed partly to the convenience of a simple ratio such as relative yield, and the easy of aggregating and comparing these ratios, but presumably also reflects their relative abundance in the literature. Searches with Google Scholar (<http://scholar.google.com>) reveal that within the mixed-species literature, there are about twenty times as many papers discussing ‘relative yield’ or ‘ANOVA’ than there are with ‘competition index’ or ‘response surface’. It is not so easy to establish why this is so. Some authors have remarked privately about the difficulty of computing spatially-explicit competition indices, but few express these views in the formal literature. Canham *et al.* (2004) mentioned that software was “written specifically for [their] study using Delphi” to enable an evaluation competition indices in Canadian hemlock-redcedar forests, a reflection on the novelty and difficulty of computing these indices. However, as this paper demonstrates, with tools like the Simile modelling environment, it need not be difficult or time consuming to compute competition indices.

The Simile Modelling Environment

Simile is a visual modelling environment available from Simulistics (www.simulistics.com) for most popular computer systems. While Simile was originally designed for simulation modelling of agroforestry systems (Muetzelfeldt and Taylor, 1997), it has been used in a wide range of other contexts (Muetzelfeldt and Massheder, 2003; Prabhu *et al.*, 2003; Vanclay *et al.*, 2003, 2006; Bonanomi *et al.*, 2005), and is equally suited to individual-based analysis of mixed-species plantings. The beauty of using Simile for such analyses is that its graphical user interface is intuitive and helps to maintain an overview even when the analysis becomes complicated. The notation used by Simile is based on the standard Systems Dynamics notation devised by Forrester (1961) and used in many modelling packages (e.g., ModelMaker, Powersim, Stella and Vensim), and is thus accessible to other researchers in the field, even those not familiar with the Simile package itself.

Using Simile to analyse data

Perhaps the best way to demonstrate the ease of individual-based analyses with Simile is to demonstrate the construction of an example. Basic Simile operations are described by Muetzelfeldt and Massheder (2003) and Vanclay (2003), and are explained in Simile's on-line help. The description below glosses over the basics, and focuses on details specific to the spatial analysis of mixed tree plantings.

Figure 1 illustrates the user interface of Simile (Version 4.4; a new version, 4.5, was released 16 January 2006). The top row of icons are the usual Microsoft utilities (new, open, save, etc.). The second row of icons are Simile's model-building tools (compartment, variable, flow, etc.). Figure 1 contains a *submodel* called *tree* created with the submodel tool () and three *variables* created with the variable tool (). It is self-evident that the construct in Figure 1 represents a single tree which is characterised by its location (X and Y coordinates only, elevation is not used), its size (currently height and diameter, but other variables such as crown width could be included), and its species (or taxon, genotype, etc.). These variables can be imported from an industry-standard CSV (comma separated variable) file, or can be entered directly via the Simile user interface.

[Figure 1 near here]

Competition studies draw on populations of trees, and this is implemented by double-clicking on the tree submodel, and indicating in the dialogue box that this is a population (note the new representation in Figure 2, with double lines reminiscent of a stack of cards). Figure 2 also includes an additional submodel labelled *pair*, and two relations (↗) labelled *tree i* and *tree j*. These two relations cause the new submodel *pair* to represent a list of every possible pairwise combination of all trees in our population. The notation (notice the dotted lines in bottom and right corners) signals that the number of instances of these pairs will depend on the number of trees and on conditions yet to be specified.

[Figure 2 near here]

The distance between each *tree* comprising a *pair* is represented using a new variable and an *influence* (↗). A single influence appears as two variable names in *pair*, because the originating variable *x,y coord* can originate from *tree i* or *tree j*. Simile generates suitable unique variable names automatically (and changes spaces and special characters to underscores; Figure 3). Figure 3 also shows the calculation of the Euclidean distance between the pairs of trees: the square root of the sum of the squares of the differences. Note that the formula involves minimal typing because functions (↗) and variables (↗) can be selected by clicking on these on the formula bar. In this demonstration, our analysis is to be restricted to pairs within a defined distance, so we make the list of pairs *conditional* (↗) by specifying, e.g., `distance<3` and `distance>0`, assuming (for the moment, following Hegyi, 1974) that the search radius will be 3 metres about each tree, and that trees do not pair with themselves (`distance=0`).

[Figure 3 near here]

Many competition indices rely on the relative size of trees; it is an easy matter to create variables representing relative size and to compute indices such as Hegyi's

(1974) index for each pair. The competition index attributed to each tree is computed as the sum for all of the pairs in which that tree participates (Figure 4). Figure 4 illustrates all that is necessary to compute Hegyi's index: a simple diagram, a condition ($\text{distance} < 3$ and $\text{distance} > 0$), and four equations embedded within the variables:

```
comp index = sum({competition_tree_j})  
competition = element([rel_size],1)/distance  
distance = (sum([x_y_coords_tree_i]-[x_y_coords_tree_j])^2)^0.5  
rel size = [size_tree_i]/[size_tree_j])
```

[Figure 4 near here]

Notice that Simile is not a “black box”, and does not dictate the computations involved: it simply provides a transparent framework to implement *my* decisions that the competition index should be the sum (cf. product) of the competition scores calculated for each pair, which are in turn based on Hegyi's concept of relative size and distance (there are many alternatives, e.g., Maily *et al.*, 2003); that distance should be Euclidean, that relative size should be based on a ratio (cf. the difference), and that the search radius should be three metres. All these assumptions are transparent and under user-control. Packages like Simile make it easy to implement these decisions on a computer, but they are no substitute for careful thought about the processes involved.

Further modifications to Figure 4 are useful. Bristow *et al.* (2006) found it useful to define the search area in terms of tree height instead of a fixed distance, and calculated a series of competition indices, each representing a partial sum for each species. This is illustrated in Figure 5.

[Figure 5 near here]

Studies involving growth over several time periods are more informative than an assessment of competition at a single instant. Figure 6 illustrates a further modification to import the data from a CSV file, to select a subset of the data pertinent to each of several occasions, and to compute the relevant competition

indices. The condition *present* tests whether a tree record is present on that occasion, and thus controls the selection of an appropriate subset of data from the data file. It also creates the possibility to examine a range of search horizons (i.e., search radii about each subject tree), with the user specifying the initial horizon and whether to use diameter or height as the basis of the competition index. Pearson's correlation coefficient (r) is calculated, to indicate how well a particular competition index and search horizon correlate with the growth rate of individual trees.

[Figure 6 near here]

At this point, it is worth reminding readers that Figure 6 is not merely a diagram of a proposed analysis; it is an image (screen-dump) of a computer-based specification for an analysis, and can readily be transformed into an executable version by pressing the *run* button (▶ in Figure 3). There is great scope in this concept of transforming the specification, e.g., into different display formats (such as html and diagrams) or into computer code compatible with various other modelling frameworks (Muetzelfeldt and Massheder, 2003). When the executable is ready, Simile will display an almost blank screen, with a number of tools for the user to create graphs, tables and other displays of any of the constructs present in the model. In the case of the model shown in Figure 6, it is useful to create a display of the horizon versus Pearson's r to find the optimal neighbourhood for defining competition (Figure 7). While Simile is used mainly as a simulation package in which time is varied, in Figures 6 and 7 there is no representation of time, and each 'time step' simply considers a new horizon. The optimum search radius is slightly ambiguous because of the many local optima (Figure 7), a result of the spacing used in the Bristow *et al.* (2006) trial. Vanclay (2006) examined ways to resolve this. When the optimum neighbourhood (and the form of the competition index) has been established, selected data (e.g., the computed competition indices) can be displayed as a table, and saved as CSV file for use in statistical packages.

[Figure 7 near here]

Figure 7 illustrates our first attempt at determining an optimum search radius for our modified Hegyi index (Bristow *et al.*, 2006). On the basis of the results illustrated in

Figure 7, we fitted a response surface (diameter increment= $f(\text{size, competition})$), with competition indices determined using a search radius equal to 1.3 times the sum of the heights of the paired trees. However, tests of our response surface revealed correlations between the parameters of this response surface and the search radius, so we iteratively revised our search radius and parameter estimates until we reached a stable optimum at $R=0.3h_{\text{euc}}+0.5h_{\text{acacia}}$ (Figure 8), where h is the tree height of the eucalypts and acacias respectively. The resulting equation is:

$$\Delta d = 0.3494 S_e + 0.1414 d - 0.0019 d^2 - 0.3948 C_{ee} + 0.0778 C_{ea} - 0.1449 C_{a^*} \quad (1)$$

(s.e. 0.1028 0.0075 0.0002 0.0363 0.0644 0.0207 $\sigma=0.594$, $df=1140$)

where Δd is the annual diameter increment, d is diameter over bark at 1.3 m above ground, S_e is a binary variable which takes the value 0 for acacias and 1 for eucalypt trees, C_{ee} is the intraspecific competition between eucalypt trees, C_{ea} is the interspecific ‘competition’ between eucalypts and acacias, and C_{a^*} is the competition experienced by acacias, irrespective of species. All the parameter estimates are significant at $P \leq 0.001$ (except C_{ea} which has $P=0.1$, one-tailed test). The first three terms in equation 1 are unremarkable, and describe a simple but realistic growth pattern with maximum growth at $d=37$ cm and a maximum attainable size of $d \approx 75$ cm, both beyond the range of the present data. Counter to expectation, the acacias seem to facilitate rather than compete with the eucalypts, as indicated by the positive parameter for C_{ea} ($P=0.1$), which implies that a eucalypt would ‘prefer’ to be adjacent to an acacia rather than an empty space (Bristow *et al.* 2006).

[Figure 8 near here]

Simile has a number of other specialised display tools that help gain insights into data, including a number of 3-dimensional representations (e.g., the ‘lollipop’ diagram in Figure 9). Despite the stylized representation of the trees, the diagram depicted in Figure 9 offers insights not evident in more conventional scatterplots, especially when the perspective view (Figure 9) is used in conjunction with other viewpoints obtained by rotating and tilting the display to show the plan and various profiles.

[Figure 9 near here]

Using Simile to explore implications

Once an equation describing the observed trends has been established with a statistics package (Vanclay, 2002), Simile can be used in a more conventional manner to examine the implications of the equation, and to examine theoretically-optimal planting designs and silvicultural regimes. To do this, the model illustrated in Figure 5 can be modified to simulate growth (the natural increase in size), instead of loading these observations from file (Figure 10). And instead of calculating Pearson's correlation (Figure 6), the standing volume of the plantation has been computed.

[Figure 10 near here]

In changing the model of Figure 5 to that of Figure 10, we have replaced the array representing tree size (dbh, ht) with a compartment (■) representing the current diameter (dbh) of each tree, a flow (→) representing the annual diameter increment (estimated using Equation 1), and a variable representing tree height (estimated through a height-diameter relationship). The model in Figure 10 also calculates tree volume (*vol*, using species-specific equations based on height and diameter), and sums these to provide an estimate of total stand *Volume*. The model now takes a set of starting conditions from file, and simulates the growth of the stand over time, allowing the user to examine the growth and production accruing to the stand with the defined starting conditions. Figure 10 deals with a single scenario (set of starting conditions), and it is somewhat tedious to examine a complete series of scenarios manually to help refine silviculture. However, it is possible to automate such a search. Figure 11 illustrates the model of Figure 10 embedded within a population of scenarios, so that Simile progressively works through the set of scenarios specified in a control file, thus allowing a systematic search for a theoretically-optimal silviculture.

[Figure 11 near here]

Figure 11 uses a simple mechanism to evaluate one scenario at a time. It employs a population to represent the set of scenarios, and employs a mechanism (Figure 11, bottom right) to ensure that only one member of this population exists at a time. When

the evaluation of a scenario is completed, that instance is removed (🗑️), and a new instance is created (🌱) with the next scenario. The model makes no attempt to employ optimization techniques (Nocedal and Wright, 2000; in part because local optima may make this difficult; see Figures 7, 8 and 13), but simply evaluates all the scenarios specified. Computationally, this may not be very efficient, but the model in Figure 11 is quick and easy to create (and modify), and reaches a solution overnight, even when a long list of scenarios (e.g., 200-300 scenarios) and an ordinary notebook computer are used, so it is adequate for an application that may not be used often, and that may be modified with every use. In such situations, ease of understanding and modifying a model may be more important than its computational efficiency.

The model illustrated in Figure 11 was used to explore some silvicultural consequences of Equation (1). In addition to Equation (1) which is inserted into the flow (📡) labelled *gro* in Figure 11, equations are also required for the height-diameter relationship. In the absence of more reliable equations, the following equations are based on limited data from the Bristow *et al.* (2006) trial:

$$h_{euc} = 1.0504 d - 0.0112 d^2$$

$$h_{acacia} = 0.8932 d - 0.0091 d^2$$

where *h* is tree height (m) and *d* is dbh (cm). A volume equation for *Eucalyptus pellita* was developed from north Queensland data (Bristow, pers. comm.):

$$v_{euc} = 0.00002663 d^2 h - 0.0252$$

where *v_{euc}* is volume (m³) to 10 cm diameter under bark. Volumes of *Acacia peregrina* were estimated using a published equation for *A. mearnsii* (Schonau, 1972):

$$\log_{10}(v_{acacia}) = 1.9532 \log_{10}d + 1.2315 \log_{10}h - 1.74059$$

where *v_{acacia}* is volume (m³) to 5 cm diameter under bark. Two further assumptions were made regarding tree growth: that trees do not shrink, so *gro* was estimated as $\max(0, \Delta d)$ where Δd is the estimate from Equation (1); and that any tree would die after three years of zero growth (a consequence of this is that in many scenarios, most acacias die before age 10, which is not inconsistent with experience).

[Figure 12 near here]

Simulations were carried out to estimate the maximum volume production at 15 years. This is not the optimum rotation age predicted by equation (1), but represents a

substantial extrapolation (i.e., 5 years) of an equation based on five years of data, and further extrapolation seemed imprudent. Simulations maximized volume simply because this avoids further assumptions about price-size relationships and discount rates. Figures 12 and 13 illustrate the optimum species composition and spacing, and way that yield responds to changes in the design. Species mix and area per tree (or stems per hectare) have the greatest effect; the placement of plants (whether species are in rows or a checkerboard pattern; whether the between- and within-row spacing is equal or rectangular) had relatively little effect on the yield. The optimal species composition is 50:50% of each species in an intimate mixture (checkerboard pattern; Figure 12), and the optimal spacing for maximum yield at age 15 is about 3.1 m square (given the assumptions about natural mortality and volume equations; Figure 13).

[Figure 13 near here]

It is not the intention of this paper to assert that a square spacing of 3.1m is optimal for mixed plantings of acacias and eucalypts, because the findings illustrated in Figures 12 and 13 are based on simplistic height-diameter relationships, on volume equations borrowed from elsewhere, and on assumptions about mortality. The intent of this paper has been to demonstrate the utility of Simile for conducting such analyses, and for exploring the implications of findings for operational forest management. The material presented in Figures 12 and 13 should not be seen as a conclusion, but as a demonstration of what can be done with trials of this kind, even when things go “wrong” (e.g., when cyclones cause unexpectedly high mortality).

Conclusions

To date, most analyses of mixed planting trials have relied on simple indices such as relative yield or on analyses of variance (ANOVA), possibly because researchers have found more insightful alternatives inaccessible. Modelling environments like Simile make it easy to gain new insights into existing data. A small investment in learning these tools pays dividends in better analyses and deeper insights.

References

- Bonanomi, G., Giannino, F., Mazzoleni, S., 2005. Negative plant-soil feedback and species coexistence. *Oikos* 111: 311-321.
- Bristow, M., Vanclay, J.K., Brooks, L., Hunt, M., 2006. Growth and species interactions of *Eucalyptus pellita* in a mixed and monoculture plantation in the humid tropics of north Queensland. *Forest Ecology and Management* 233:285-294.
- Connolly, J., Wayne, P., Bazzaz, F.A., 2001. Interspecific competition in plants: how well do current methods answer fundamental questions? *American Naturalist* 57: 107–125.
- Forrester, J.W., 1961. *Industrial Dynamics*. MIT Press.
- Forrester, D.I., J. Bauhus, A. Cowie and J.K. Vanclay, 2006. Mixed-species plantations of Eucalyptus with nitrogen fixing trees. *Forest Ecology and Management* 233:211-230.
- Grimm, V., 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling* 115: 129–148.
- Grimm, V., Railsback, S.F., 2005. *Individual-based Modeling and Ecology*. Princeton University Press, 480 pp.
- Hegy, F., 1974. A simulation model for managing jack-pine stands. *Growth Models for Tree and Stand Simulation* (ed. J. Fries), pp. 74–90. Royal College of Forestry, Stockholm.
- Huston, M., DeAngelis, D., Post, W., 1988. New computer models unify ecological theory. *BioScience* 38: 682–691.
- Mailly, D., Turbis, S., Pothier, D., 2003. Predicting basal area increment in a spatially explicit, individual tree model: a test of competition measures with black spruce. *Can. J. For. Res.* 33: 435–443.
- McIntosh, B.S., Jeffrey, P., Lemon, M., Winder, N., 2005. On the Design of Computer-Based Models for Integrated Environmental Science *Environmental Management* 35: 741–752.
- Muetzelfeldt, R., Massheder, J., 2003. The Simile visual modelling environment. *Europ. J. Agronomy* 18:345-358.
- Muetzelfeldt, R.I., Taylor, J., 1997. The suitability of AME for agroforestry modelling. *Agroforestry Forum* 8 (2): 7-9.
- Nocedal, J., Wright, S.J., 2000. *Numerical Optimization*. Springer, 636 pp.
- Park, S.E., Benjamin, L.R., Watkinson, A.R., 2003, The theory and application of plant competition models: an agronomic perspective. *Annals of Botany* 92:741–748.
- Prabhu, R., Haggith, M., Mudavanhu, H., Muetzelfeldt, R., Standa-Gunda, W., Vanclay, J.K., 2003. ZimFlores: A model to advise co-management of the Mafungautsi Forest in Zimbabwe. *Small-scale Forest Management, Economics and Policy*, 2(2):185-210.
- Schonau, A.P.G., 1972. In Anonymous, 1985. *Acacia mearnsii*: Multipurpose Highland Legume Tree. NFT Highlights, NFTA 85-02, May 1985.
http://www.winrock.org/forestry/factpub/FACTSH/A_mearnsii.html [5 Jan 2006].
- Vanclay, J.K., 2002. A personal view of statistical packages for linear regression. In: S. Harrison, J. Herbohn, E. Mangaoang and J. Vanclay (eds) *Socio-economic Research Methods in Forestry: A Training Manual*. Rainforest CRC, Cairns, ISBN 0 86443 691 2, pp. 65-72.

- Vanclay, J.K., 2003. The One-minute Modeller: An Introduction to Simile. *Annals of Tropical Research*, **25**(1):31-44.
- Vanclay, J.K. 2006. Experiment designs to evaluate inter- and intra-specific interactions in mixed plantings of forest trees. *Forest Ecology and Management* 233:366-374.
- Vanclay, J.K., Sinclair, F.L., Prabhu, R., 2003. Modelling interactions amongst people and forest resources at the landscape scale. *Small-scale Forest Management, Economics and Policy*, **2**(2):117-120.
- Vanclay, J.K., Prabhu, R., Sinclair, F.L., 2006. *Realizing Community Futures: A practical guide to harnessing natural resources*. Earthscan, London, 176 pp, ISBN 1 84407 383 1.

Figures

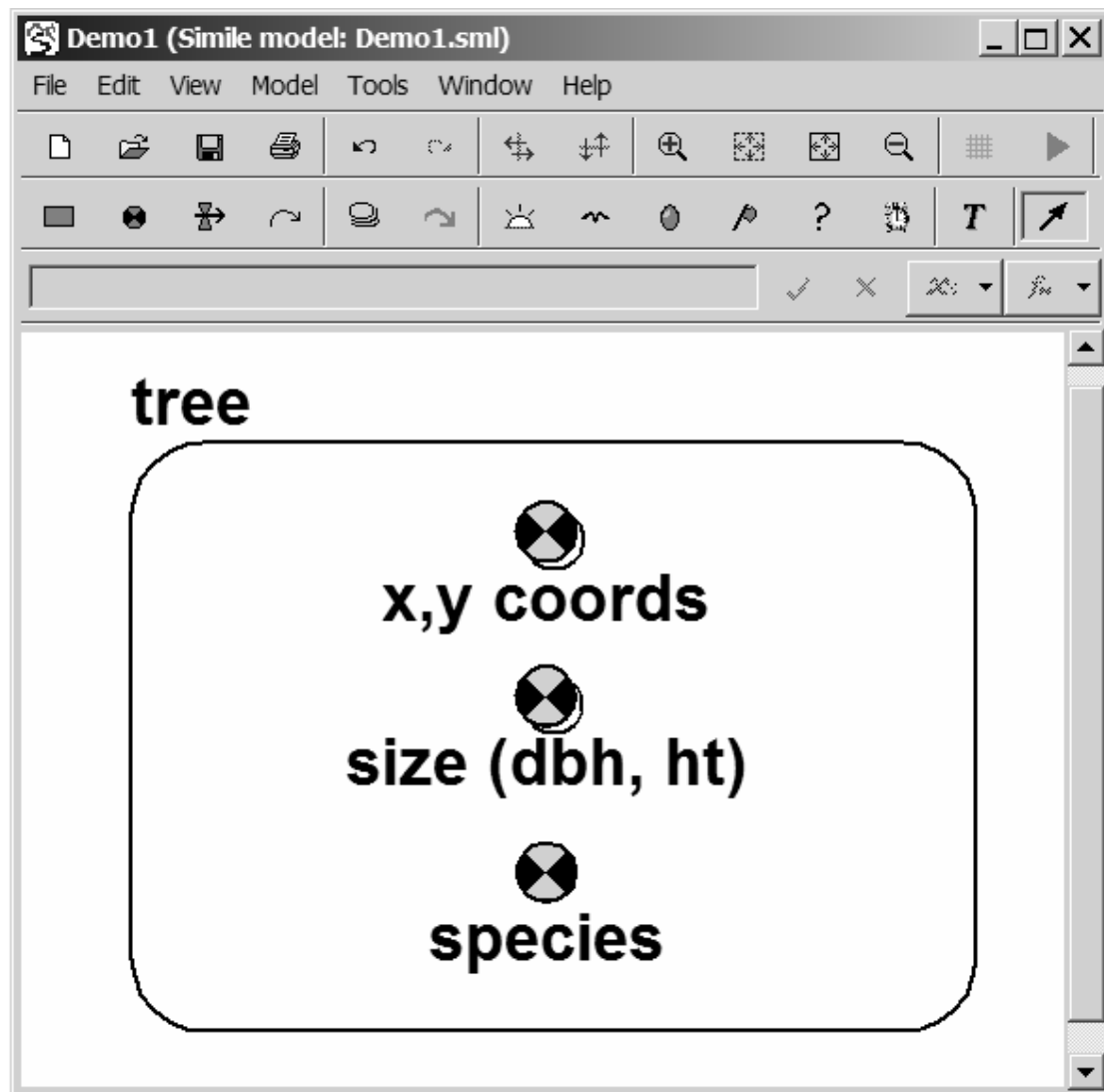


Figure 1. The graphical user interface of the Simile modelling environment, with a representation of several attributes of an individual tree.

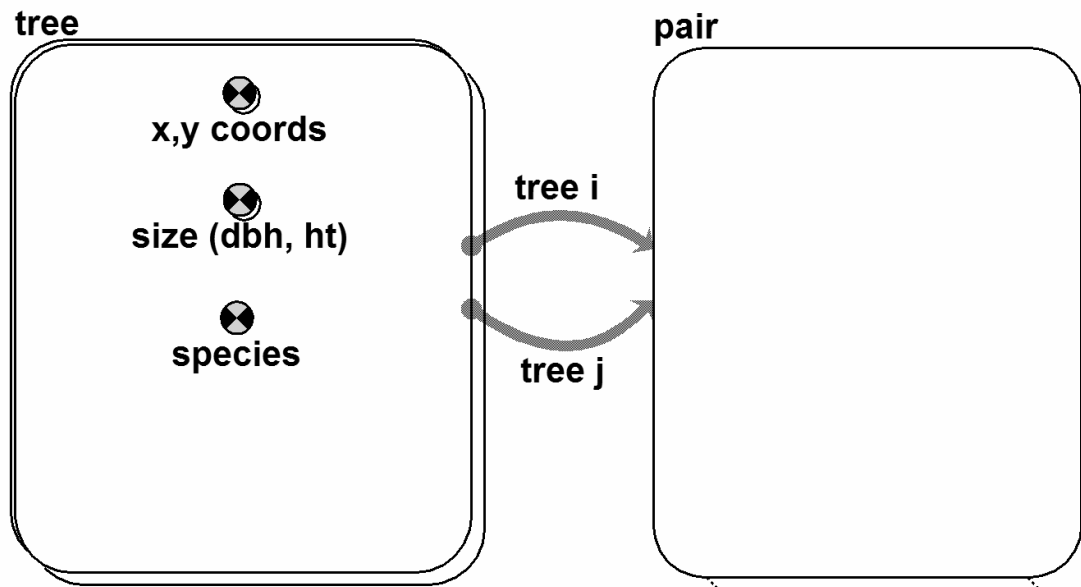


Figure 2. Simile representation of the relations defining pairwise interactions between individual trees in a population.

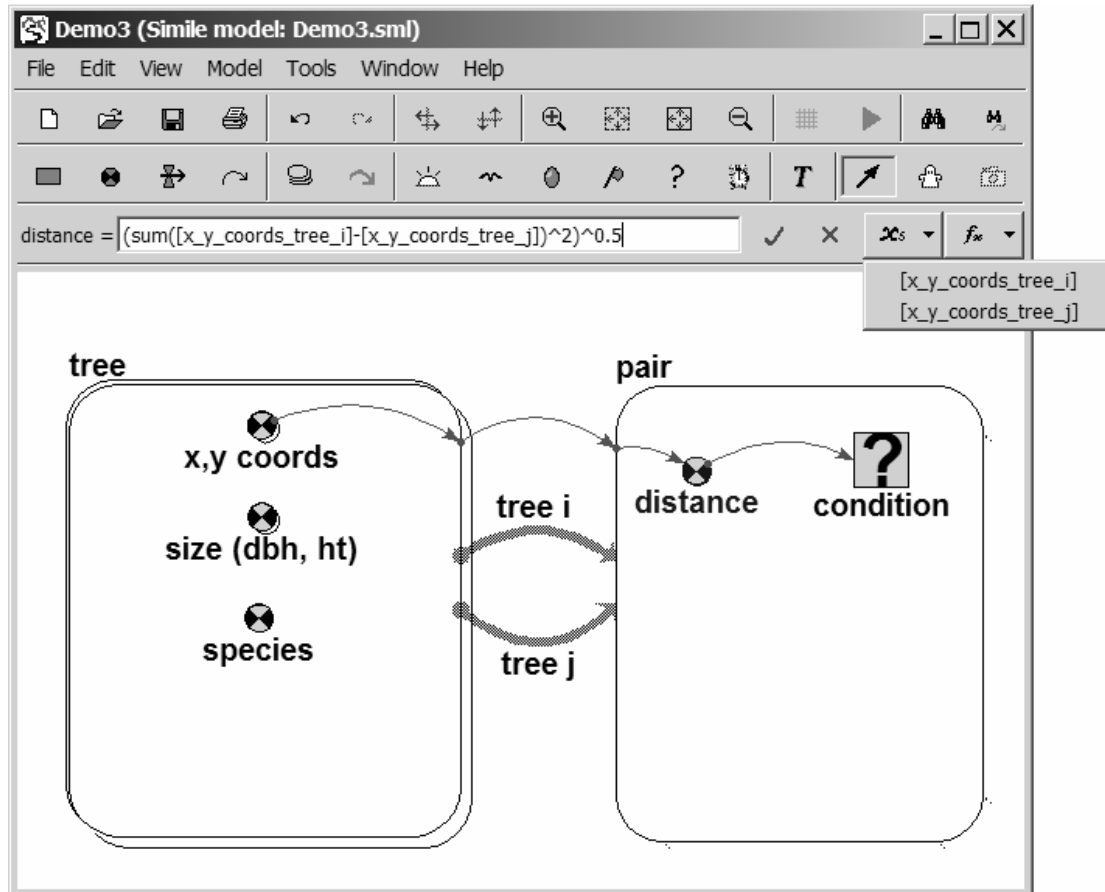


Figure 3. Simile diagram with a conditional submodel representing the interaction between individuals within a defined Euclidean distance.

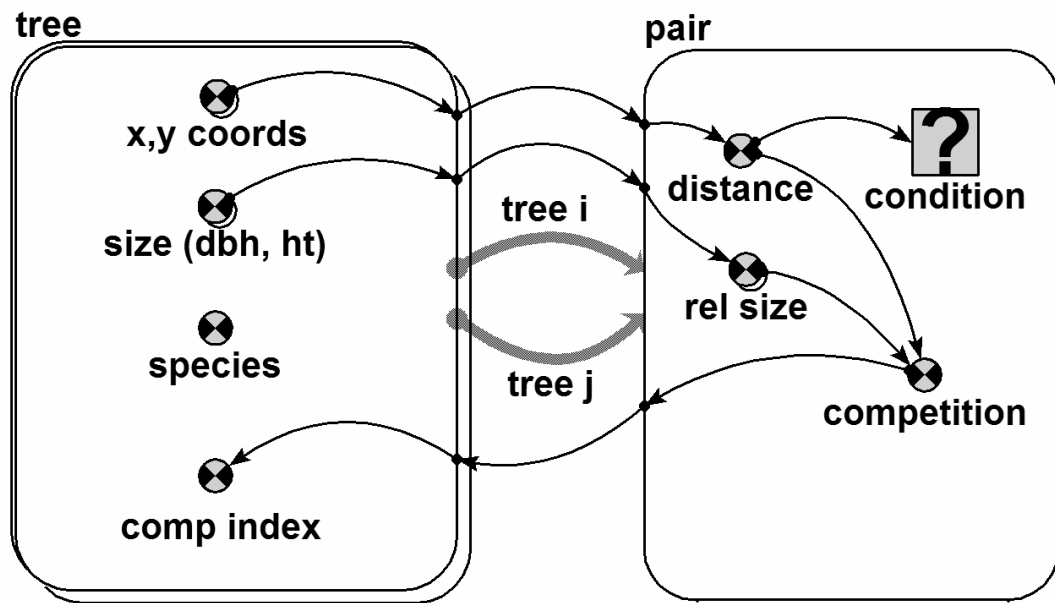


Figure 4. Simile representation of the calculation of Hegyi's (1974) competition between pairs of individual trees.

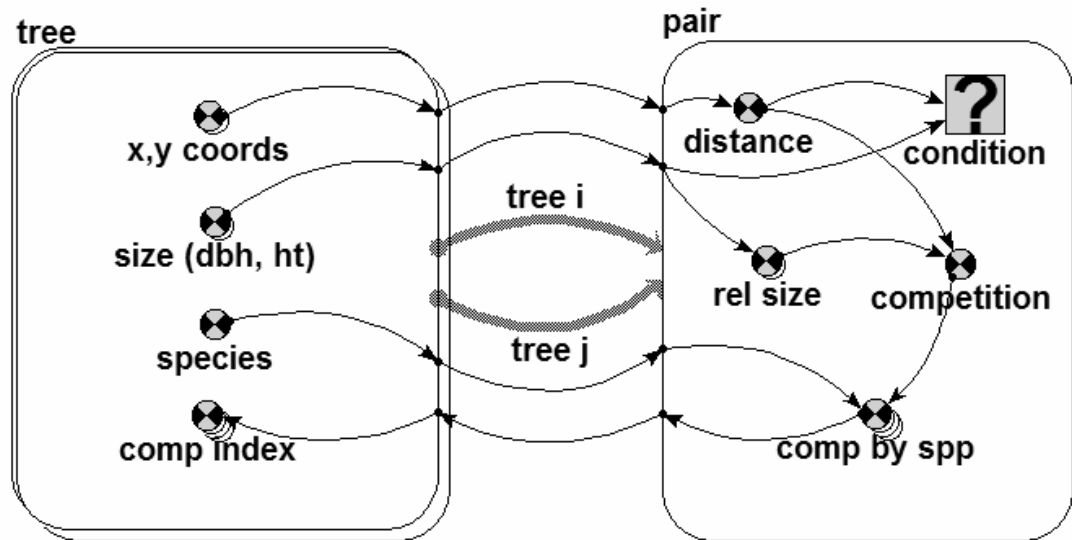


Figure 5. Simile representation of the Bristow *et al.* (2006) competition index.

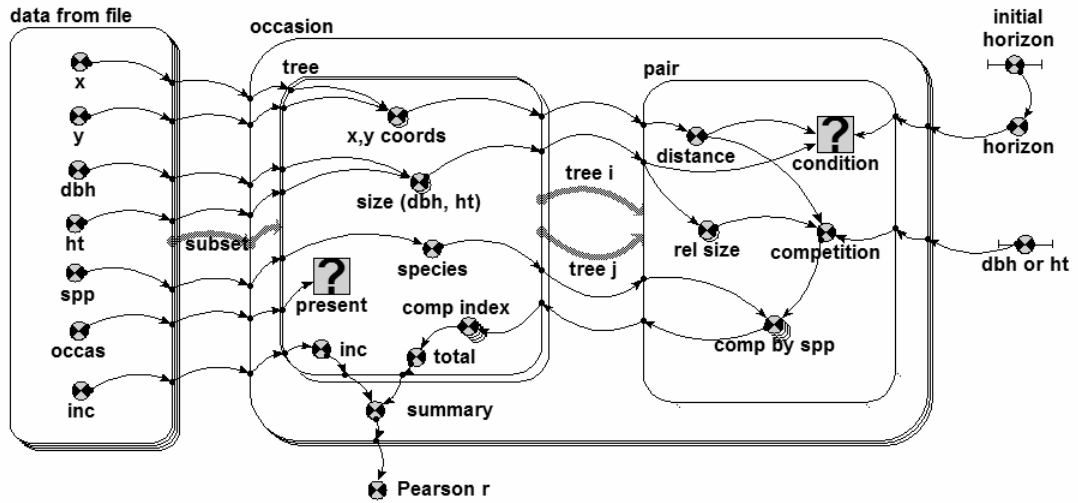


Figure 6. Simile representation of the calculation of competition indices for a data file with data drawn from several occasions.

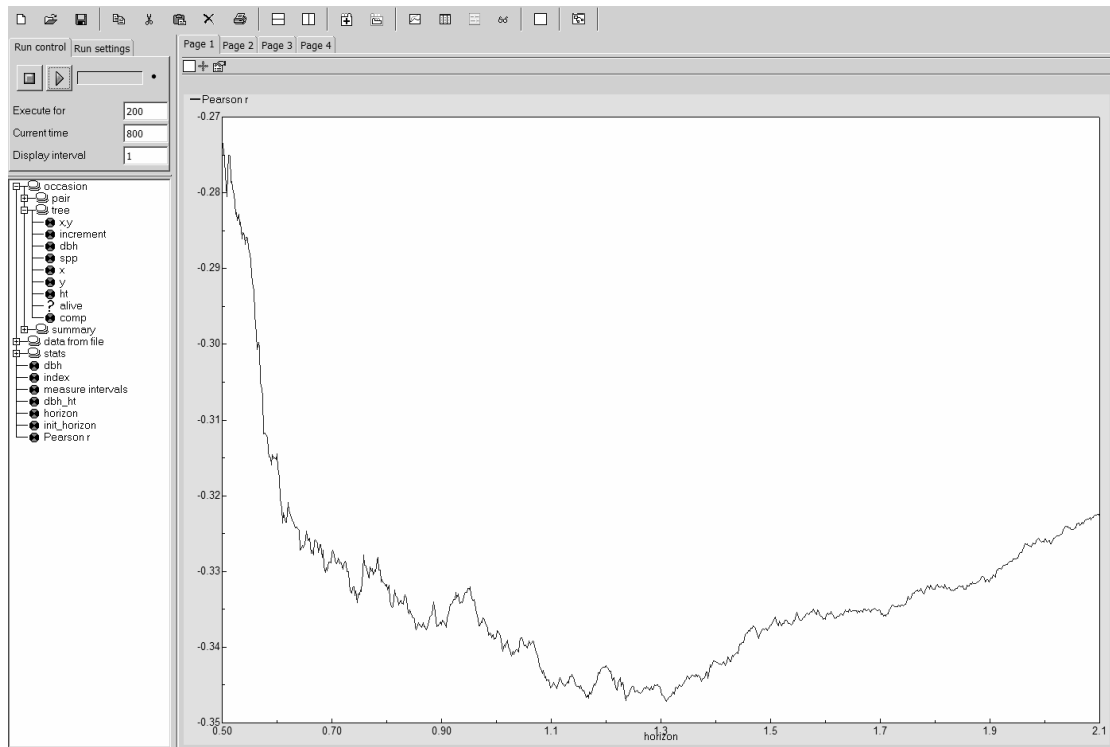


Figure 7. Pearson's correlation coefficient (r) between relative increment (increment/diameter) and search radius, calculated for the data of Bristow *et al.* (2006). The strongest correlation occurs at about 1.3 times the sum of the tree heights.

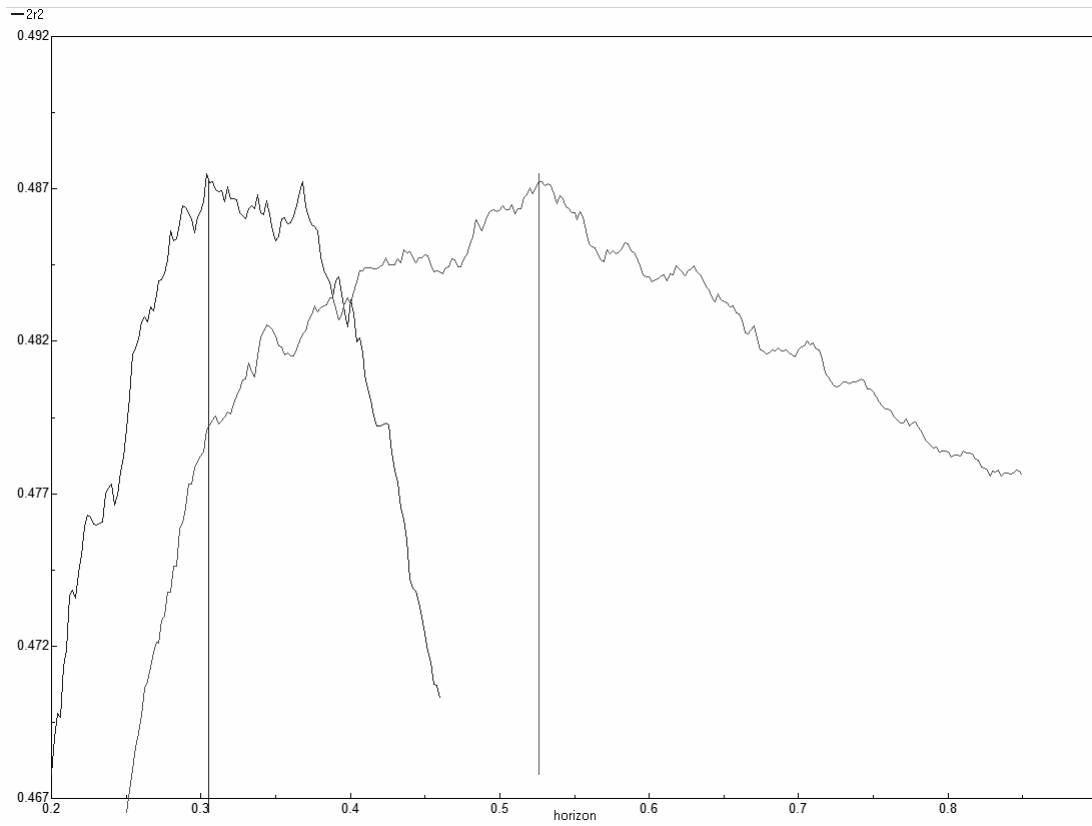


Figure 8. The correlation between observed and predicted diameter increment using competition indices based on range of search radii for both eucalypts (left peak) and acacias (right peak). The optimum search radius is $0.31h_{euc}+0.52h_{acacia}$.

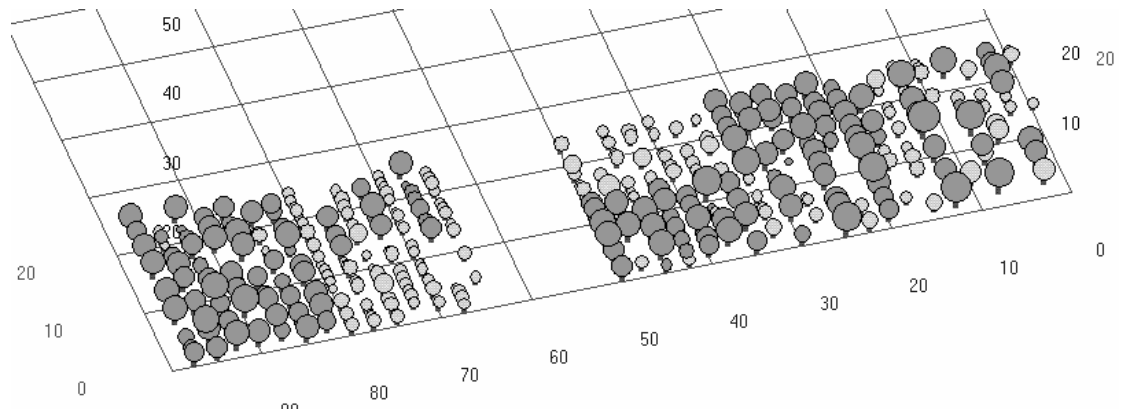


Figure 9. Simile's 'lollipop' diagram of the Bristow *et al.* (2006) mixed-species trial at age 10. Dark grey symbols represent eucalypts and light grey represents acacias. Note that all distances have been divided by 2 to adapt to 100×100 unit display used by the 'lollipop' diagram. Crown sizes and shapes are schematic, and not necessarily accurate.

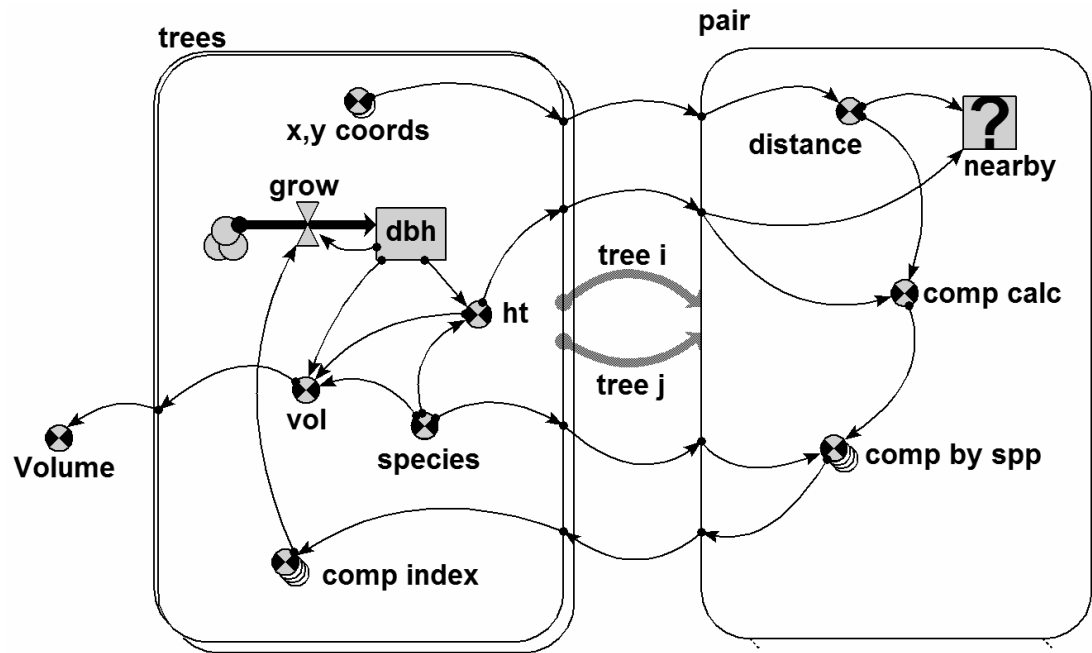


Figure 10. Simile representation of a simulation model to examine future growth of a mixed planting.

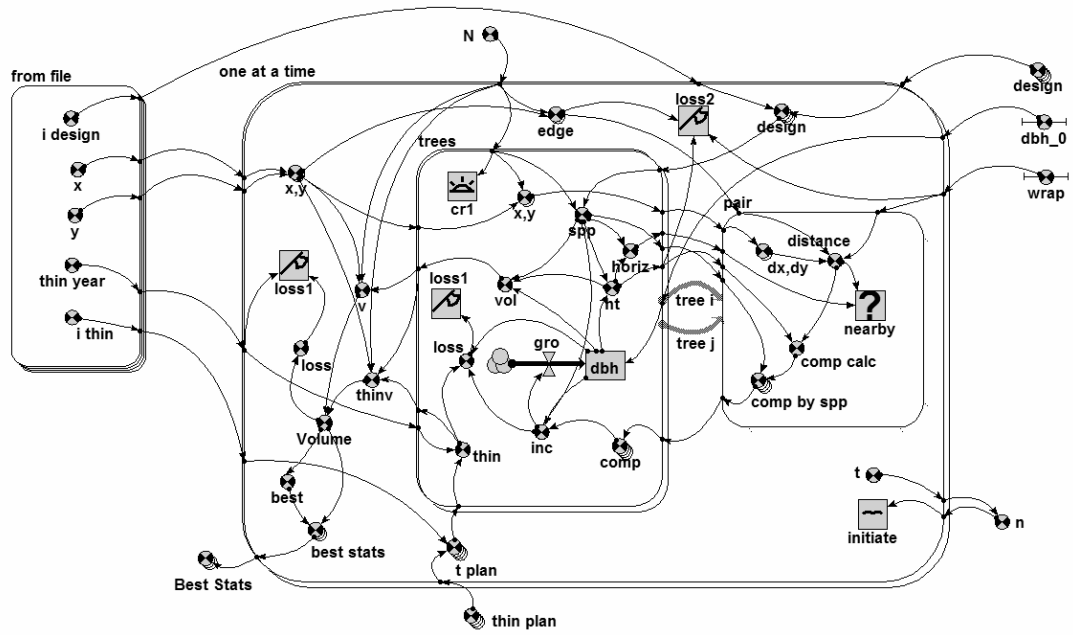


Figure 11. Simile representation of a model to examine a large number of scenarios, one at a time.

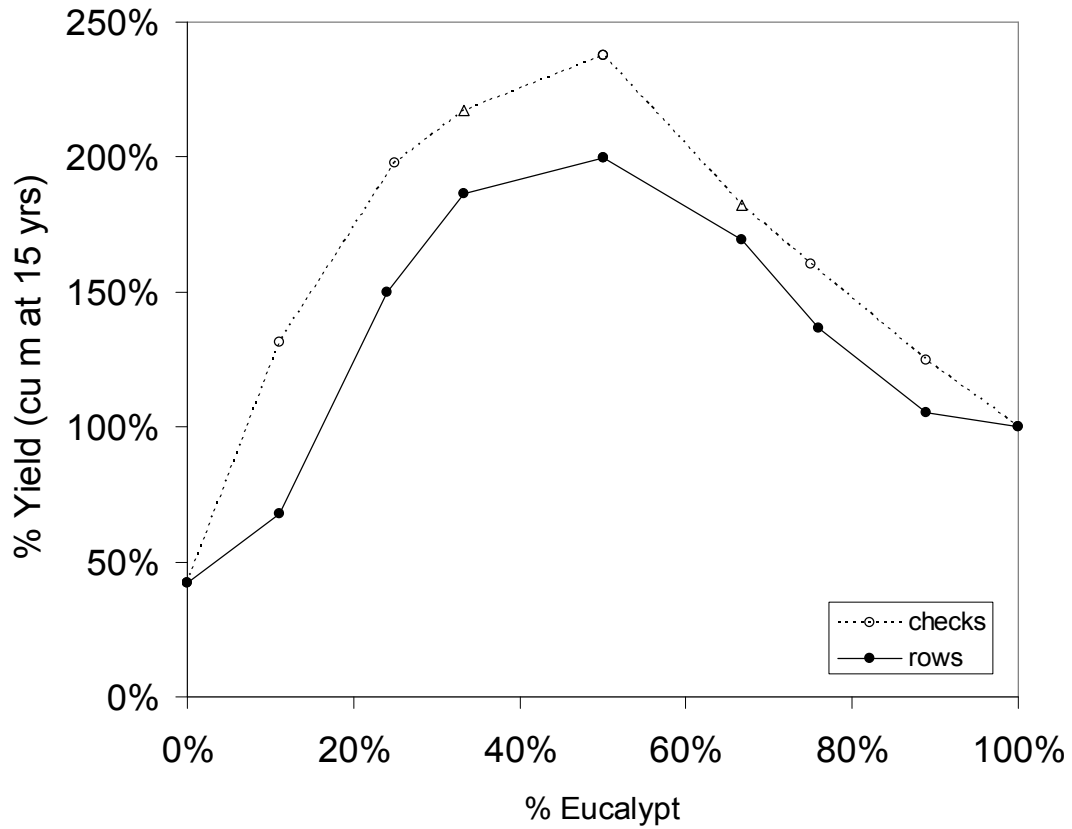


Figure 12. Simulated yields (total stand volume expressed as a percentage of the eucalypt yield) of various mixtures of eucalypts and acacias. Intimate mixtures (i.e., checkerboard patterns) offer higher yields (dotted line) than planting arrangements with monospecific rows (i.e., striped patterns). Note that 33:67 and 67:33 mixtures do not form checkerboard patterns, but compare diagonal stripes (Δ) with horizontal and vertical strips (\bullet).

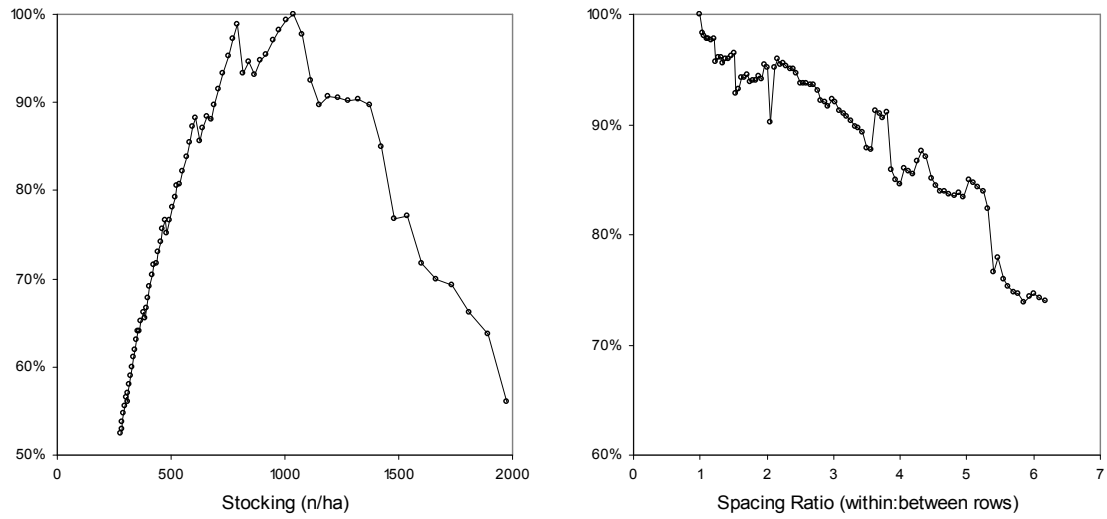


Figure 13. Simulated results illustrating the effect of spacing of the yield of a mixed stand with an equal (checkerboard) mix of acacia and eucalyptus. The left figure represents a square spacing ranging from $2 \times 2\text{m}$ to $6 \times 6\text{m}$ (cf. optimum at $3.1 \times 3.1\text{m}$). The right figure shows the effect of rectangular spacing ranging from $3.1 \times 3.1\text{m}$ to $7.7 \times 1.25\text{m}$ (while maintaining constant stocking).