Function, form, environment and ecological interactions: what's the right way to model a plant?

M. Renton [a,b]

[a] *School of Plant Biology, UWA Institute of Agriculture, University of Western Australia*
[b] *CSIRO Ecosystem Sciences*
Email: michael.renton@uwa.edu.au

Abstract: A plant consists of components such as leaves, roots, stem, flowers and buds. Over time the form of the plant changes; new parts grow, existing parts change their shape or size and old parts fall off or are removed. From this perspective, plants can be seen as dynamic structures of connected parts. Intimately related to this changing of the external form and size of the plant are processes happening inside the plant. From this perspective, plants can be seen as biochemical machines; using the energy of the sun to assimilate carbon and converting this carbon into the material of which they are made; obtaining and distributing other nutrients and water from the soil. Plants are also continually responding to and affecting their environment and other organisms. Plant growth thus involves the interaction between this structural development, these physiological processes, the physical environment that surrounds them, and the animals, other plants, fungi and bacteria with which they interact. This four-way interaction results in plant behaviour that is variable, adaptable, complex and difficult to fully understand and predict.

Simulation modelling provides a tool that can help understand and predict the behaviour of complex systems, such as plants and plant ecosystems, by synthesising available knowledge and hypotheses. But how should a plant be modelled? Traditionally the dynamics of populations of plants or the growth of individual plants was simulated with differential equations that integrated processes across time and space. Increases in computational power allowed these models to be extended to include much more detailed and realistic representation of the spatial and temporal variability of the plant's environment, such as variation in moisture levels through a soil profile, light through a canopy, and rainfall through a season. Computational power also opened up the possibility of individual-based modelling (IBM) approaches, which can account for the variability between individual plants by representing each plant separately, together with the interactions between them. The functional-structural plant model (FSPM) approach goes further still by separately representing the individual components of the plant, such as leaves, internodes, flowers or even cells, together with their interactions. These approaches open up the possibility of modelling at greater and greater levels of detail and realism, but does modelling at greater levels of detail and realism necessarily help a model achieve its aim of helping understand and predict the behaviour of a plant or a plant ecosystem?

In this paper, I present a range of examples of FSPMs and IBMs involving plants and their ecosystems. I show how these relatively new computational approaches can be used to address a range of purposes across a spectrum of spatial and temporal scales, from the movement of molecules of plant hormones in individual plant cells during a fraction of a second, to the migration of plants across landscapes under changing climates during hundreds of years. I point out the advantages of FSPM/IBM approaches in each of these examples, while also arguing that choices about the amount of detail to be included in a FSPM/IBM plant model should always be made with the objective of achieving the purpose of the modelling study as effectively but efficiently as possible.

Keywords: *Functional Structural Plant Model, Individual Based Model, PERTH, Weed Seed Wizard, PPUNCC, CCCEPAE, auxin, MAppleT, birch, seagrass*

1. INTRODUCTION

Individual-based models (IBMs) of plants are computational simulations that represent the state of a number of individual plants over time, and how the state of each plant depends on interactions between these plants and their environment (Berger et al., 2008, Grimm and Railsback, 2004, Judson, 1994). The IBM also simulates the broader-scale result that emerges from these individual-scale processes. As well as the individuals, IBMs typically include a representation of an 'environment' in which individuals are situated and interactions occur. IBMs can be contrasted with models based on pools of individuals, and the size and characteristics of these pools are tracked over time. In an IBM, every individual can have different values for its characteristics, while in a pooled model, individuals within a pool are assumed to have the same values for their characteristics. IBMs may (or may not) be spatially-explicit, meaning that the state of the individual plants includes specification of their location in geometrical space.

The functional-structural plant model (FSPM) approach goes even further than the IBM approach by separately representing the individual components of the plant, such as leaves, internodes, flowers or even cells and particles of plant hormones within these cells (Fourcaud, et al., 2008, Godin and Sinoquet, 2005, Hanan and Prusinkiewicz, 2008, Vos et al., 2007). They then track the interactions and relationships between these components and their environment. FSPMs can be thought of as IBMs where the 'individuals' are the plant components rather than the whole individual plants. FSPMs can be contrasted with models based on pools of plant components, such as crop growth models that track the number of leaves within different canopy layers. FSPMs are generally both topologically- and spatially-explicit, meaning that each component has a specified location and orientation in geometrical space, and a specified set of other components that are connected to it. Interaction between components may occur through the external environment based on their relationship in geometrical space (eg. shading, rain-splash) or through their topological connection (eg. hormonal signalling, movement of water).

In the next section, I briefly present a range of examples of IBMs and FSPMs involving plants and their ecosystems, to show how these relatively new computational approaches can be used to address a range of purposes across a spectrum of spatial and temporal scales. For each example, I describe the purpose or motivation for constructing the model. I also give a brief overview of the model itself, with a focus on how and why the IBM or FSPM approach has been used, including the spatial and temporal resolution of the representation of plant and its environment, and also the extent to which the model is mechanistic versus descriptive. The list of models is no way meant to be comprehensive, but rather provide a set of contrasting examples to illustrate the uses and advantages of these computational approaches.

2. EXAMPLES

2.1. Understanding and predicting evolution of resistance to herbicides

The Polygenic Evolution of Resistance To Herbicides (PERTH) model enables analysis and prediction of the effect of different long-term management options on the risk of evolving herbicide resistance in agricultural systems, while also accounting for the effects of weed biology and ecology, population dynamics, and the genetics underlying resistance. Implemented in the Python programming language, the model simulates the dynamics of weed population numbers and genetics over many years, together with their effects on crop yield (Renton et al., 2011b). 'Individuals' are plants or seeds, and for each individual, the model tracks its genotype (which does not change with time) whether it is currently a seed or a plant, its age (if a plant) and its depth in the soil (if a seed). The effect of events like germination, herbicide application, reproduction and tillage are simulated individually and stochastically. The effect of herbicide application, for example, depends on application rate and timing, but also on the genotype of the individual. Reproduction is also handled individually and stochastically, with the source of pollen for a new seed chosen at random from all potential plants, and then new the individual alleles at each locus chosen at random, one from each parent. The model works on an annual time step, with several sub-steps including germination at different points of the year, herbicide applications, and crop sowing and harvesting. The only spatial aspect of the model is that seeds are located in one of several soil layers (which can be changed by simulated soil cultivation). Processes such as mortality due to herbicide and the determination of the genotype of the seed produced are handled individually at a relatively high level of biological realism while competition with crop and the number of seeds produced is simulated using a relatively simple and descriptive model with all plants in a single pool.

2.2. A practical decision-aid tool to help farmers manage weed populations

The Weed Seed Wizard was developed with the aim of creating a practical decision-aid tool that can help farmers and consultants manage weed populations in real agricultural contexts. Implemented in the Java programming language, the model uses detailed and specific paddock management and weather records, and simulation of important aspects of seed biology, in order to track and predict the number, ages, soil depth, dormancy levels, viability and germination of seeds in the soil, to predict the amount of weeds appearing each year (Renton et al., 2008). In the original version of the model, seeds were represented individually, so that full details of the soil depth over time and the resulting environmental history (temperature, exposure to moisture, etc) of every seed could be tracked. Germination depends mechanistically on this history, with certain requirements for cumulative thermal time and moisture. Other processes within the model, such as movement within the soil due to tillage or crop sowing, mortality of plants due to herbicides, and competition between weeds and crop, are represented relatively empirically. The spatial resolution of the soil environment is continuous in the vertical dimension, with functions giving the temperature and moisture at any depth. There is no representation of space in the horizontal dimensions. The model runs on a daily time-step.

2.3. Predicting persistence of plant species under climate change in fragmented landscapes

The purpose of Plant Persistence UNder Climate Change (PPUNCC) model was to enable prediction of whether a particular plant species would be able to spread fast enough across a fragmented landscape to 'keep pace' with a changing climate and thus survive, when the area of their original distribution becomes climatically unsuitable (Figure 1, Renton et al., 2011c). Plant species are characterised in terms of functional traits such as the time it takes to reach maturity, the amount of seed it produces, its probability of establishing in different climates, and the distances its seeds are likely to disperse, and then the model accounts for these values in representing processes such as maturation, seed production and dispersal. Aspects of the landscape such as degree and pattern of fragmentation, and the speed of climate change are also accounted for. It is implemented in the R software environment. 'Individuals' are plants or seeds. For each individual plant, the model tracks its location in the landscape, and its age. When seed is produced, it is immediately dispersed across the landscape. Dispersal, establishment, and mortality are simulated individually and stochastically. The model works on an annual time step. The model is spatially explicit, with the landscape represented as a grid of square cells, each of which may or may not be suitable for the plant species, and each plant associated with a particular cell. The processes within the model such as dispersal, establishment, and mortality are represented using relatively simple and descriptive models.
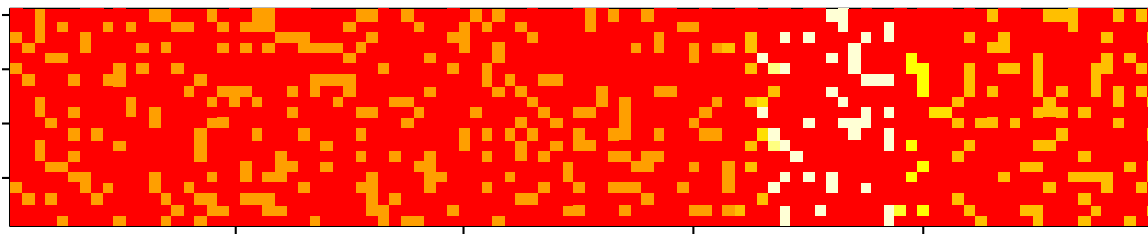

Figure 1. The PPUNCC model simulates the population dynamics and migration of a plant species across a fragmented landscape under a changing climate. Red cells are unsuitable for the species due to land-use or soil-type. The colour of other cells indicates population density (white=high density, orange=unoccupied). The figure represents a snapshot in time as the plant's suitable climate shifts from right to left.

2.4. Competition and coexistence in communities of ephemeral plants in arid environments

The purpose of this model (abbreviated as CCCEPAE) was to introduce ecology students to computational modelling and to help them explore tradeoffs in ecological strategies. The simulated system represents communities of ephemeral plants in an arid environment with occasional rain (Figure 2). A plant species is represented in terms of just a few characteristics: the mean dispersal distance of its seeds, the chance of survival of its seeds at each time step, the threshold soil moisture level required to trigger germination, and the maximum plant size before seed production and senescence occurs. It is assumed that there is a trade-off between seed dispersal distance, seed longevity, and total number of seed produced. Plants that produce 'higher quality' seeds that have better dispersal mechanisms or survival mechanisms produce less quantity of seeds in total. The CCCEPAE model simulates plant growth and water use over time, senescence and seed

production. Individuals of the same and different species compete for water over many generations. Through a guided computer laboratory class, students manipulate the number of species simulated, the starting number of each species, the parameters characterising the species, and aspects of the environment such as soil water holding capacity and rain frequency and regularity. They explore the effects of these manipulations on species richness and composition over time and spaces, and whether the system ultimately tends towards stable coexistence or competitive exclusion and dominance by a single species. 'Individuals' are plants or seeds. For each individual plant, the model tracks its location in the simulated area, and its size. When seed is produced, it is immediately dispersed across the area. Dispersal, germination, and mortality are simulated individually and stochastically. The model works on a daily time step. The model is spatially explicit, with the area represented as a grid of square cells, each with a specified soil water level, which changes dynamically with plant water use and evaporation. Each plant has a specified location, and is associated with a particular cell. Plant growth and water use is represented relatively mechanistically, while other processes within CCCEPAE such as dispersal, germination, and mortality are represented more empirically.
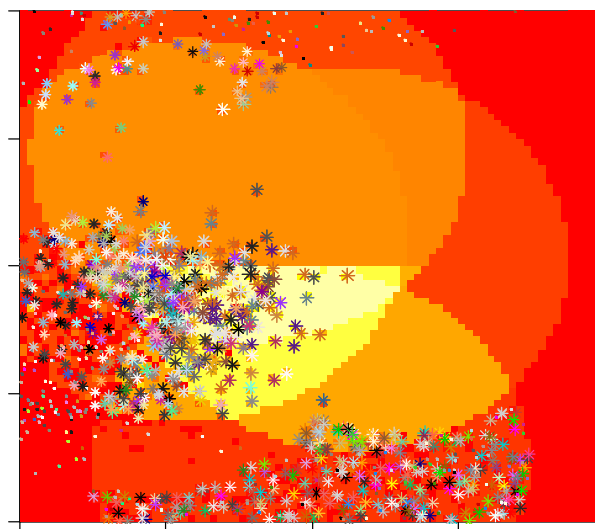
Figure 2. Simulated communities of ephemeral plants (stars) and dormant seeds (points) of different species (different colours) in arid environments with occasional patchy rain affecting soil moisture level (red→white representing dry→wet).

2.5. Understanding cellular-scale processes of auxin transport through a plant stem

This model was constructed to help understand how patterns of auxin transport through a plant cell, and signalling dependent on auxin levels, depends on cellular-scale processes (Renton et al., 2004). It is based on current understanding of these processes. The model represents the stem as a series of connected cells, and auxin as discrete (individual) particles moving within and between cells. Within the cell, auxin particles move by diffusion, but particles cannot move beyond the boundaries of the cell without the assistance of special 'transporters', that move within certain proportions of the cell, and can only transport one particle at a time. By varying parameter values, we could explore the effect of different mechanistic assumptions regarding aspects such as the rate of diffusion within a cell, the time for transport between cells, and the relative number of transporters per cell. The temporal resolution of the model is variable, but in the order of fractions of a second. The spatial resolution is also variable, but in the order of hundredths of a millimetre.

2.6. Birch tree structure developing from local-scale processes

The FSPM L-system-based birch model was constructed to help understand how the full structure of a particular species of birch tree emerges over many years dependent on the influence of local-scale factors on whether buds produced new growth units in any particular year (Figure 3, Renton et al., 2003, Renton et al., 2005). The rules governing bud fate are functions that give the probability of producing a new growth unit, depending on local-scale factors (mainly light levels), and are based on extensive empirical studies that took a range of possible explanatory environmental factors into account. The time-step of the model is bi-annual, although it could also be considered as annual, since growth processes only occur every second step, and the intermediate step is included for visualisation only. The tree is represented as a collection of interconnected plant components (nodes, internodes and leaves) in the typical FSPM fashion.

2.7. Apple tree structure and fruit position

The FSPM L-system-based MAppleT model was constructed to help predict how the full structure of an apple tree emerges over several years dependent on which buds produce new growth units of particular types in any particular year (Figure 3, Renton et al., 2006, Costes et al., 2008). The rules governing eventual bud fate are purely empirical statistical models (hidden semi-Markov chains), and are based on extensive empirical measurements. These did not take possible explanatory environmental factors into account, but

were purely based on the position of the bud within the tree and growth unit, and are thus much more purely descriptive and less representative of mechanism than those for the birch model. The time-step of the model is the plastochron, which is approximately weekly. The tree is represented as a collection of interconnected plant components (nodes, internodes and leaves) in the typical FSPM fashion.
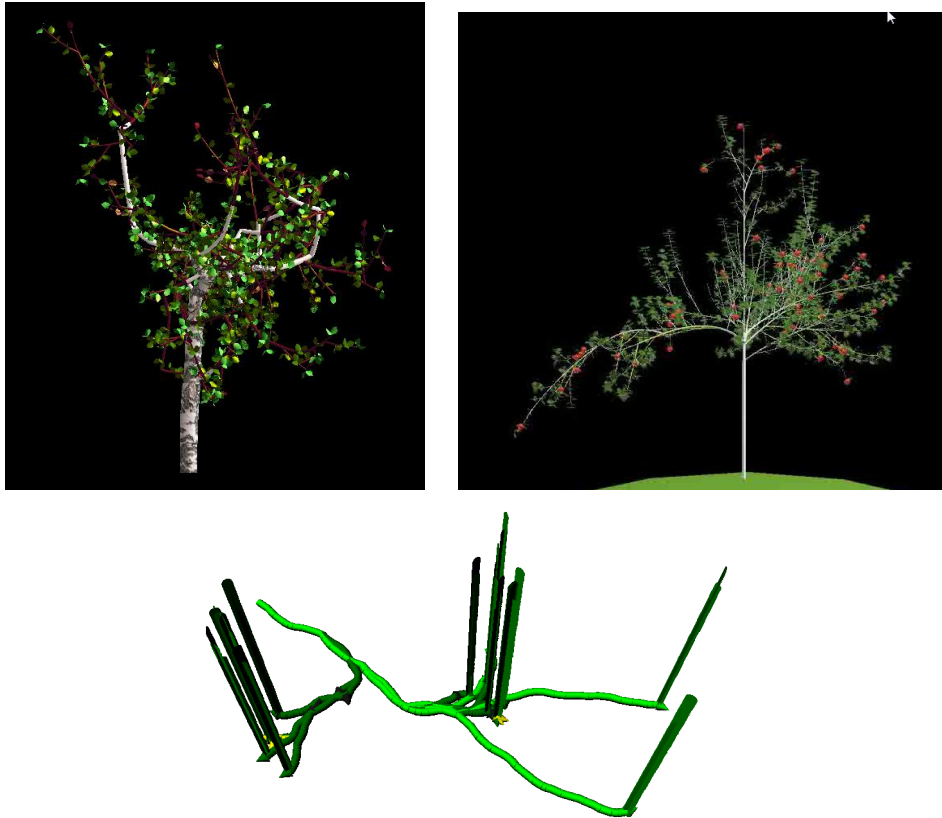
Figure 3. Example outputs from the birch (top left), apple (top right) and seagrass (bottom) models.

2.8. Predicting patterns of seagrass growth to identify optimal transplant strategies

The FSPM L-system-based seagrass model was constructed to help predict how different transplanting patterns of seagrass rhizomes would affect the subsequent growth patterns of the seagrass across the ocean floor, and thus identify optimal transplanting strategies to achieve desired outcomes in seagrass restoration programs (Figure 3, Renton et al., 2011a). The time-step of the model is the plastochron, which is monthly. The plant is represented as a collection of interconnected plant components (nodes, internodes and leaves) in the typical FSPM fashion. The full structure of a plant emerges over several years dependent on various functions and rules. Lengths of internodes, numbers of nodes between branching points, and branching angles are generated stochastically from distributions summarising extensive empirical measurements. The rules governing growth rates and probabilities of mortality and branching for a given apical bud are influenced by the 'state' of the bud, which depends on the number of surrounding buds. The parameters for these rules were based on knowledge of the seagrass species, and were calibrated empirically. The rules did not take environmental factors besides surrounding bud density into account, so are thus more purely descriptive and less representative of mechanism than those for the birch model, but more mechanistic than those for the apple model.

3. DISCUSSION AND CONCLUSIONS

The examples summarised in this paper illustrate a range of applications of the IBM and FSPM approaches. They also illustrate how different design choices were made in constructing the models, regarding aspects such as the spatial and temporal resolution of the representation of the plant(s) and their environment, and the extent to which the model is explanatory, with mechanism included in the model, versus the model being purely descriptive.

The base spatial resolution of the representation of the plant varied from the level of individual cells in the intra-cell auxin model, to the level of individual plants in the IBMs such as the Weed Seed Wizard, PERTH, CCCEPAE and PPUNCC. These models also included various other aspects of the plant, such as an individual representation of the auxin particles in the intra-cell model and the various possible alleles in the PERTH model. The FSPM models were intermediate, with plants represented as a spatially and topologically related set of plant components. In the birch, apple and seagrass models, the basic units were internodes, nodes and leaves, with the apple also including flowers and fruits. The temporal resolution of the model also varied widely, from an annual time step in the PERTH and PUNCC models, bi-annual in the birch model, monthly in the seagrass model, approximately weekly in the MAppleT, daily in CCCEPAE and the Weed Seed Wizard and fractions of seconds in the auxin model. It can be noted that a higher spatial resolution does not always correspond to a higher temporal resolution.

The representation of the environment also varied widely between models. This is closely related to degree to which the model is empirical/descriptive versus mechanistic/explanatory. In the apple and the seagrass FSPMs, there is no real representation of the environment at all, since these two models are relatively empirical and descriptive, based on data collected from plants in particular environments. Nonetheless, both models were designed to be easily adapted to new locations and situations wherever empirical data is available, and plans for future work for both include incorporating deeper representations of mechanisms. In the birch model, the only aspect of the environment that is represented is light, since the previous analysis showed that this was the most important factor controlling which buds produced new growth. Even though this model was based on empirical data (on bud fate), the data was collected across a range of environments. This allowed us to fit empirical models predicting each bud's fate from its environment. The overall birch model is thus relatively mechanistic, since the developing tree structure emerges based on these underlying rules governing bud fate, rather than being purely descriptive. It can be noted that models with the same spatial resolution (the birch, apple and seagrass models) can range from relatively descriptive to relatively mechanistic.

In the auxin model, the only representation of the environment external to the plant that is represented in the model is the application of exogenous auxin. And our analyses showed that the pattern of input of endogenous auxin had little influence on the speed at which the auxin wave moved down the stem, or on signal pattern and speed, so in this sense, plant environment was of little import. However, the 'individuals' in this model could be seen as the particles of auxin, and the 'environment' for these would then be the cells and perhaps the transporters. In this sense, the 'environment' had a very important influence on the speed and pattern of auxin movement and hormonal signalling. Certainly the model is mechanistic, with the patterns of auxin movement and signalling 'explained' by the represented mechanisms of auxin transport included in the model, which depends on the interaction between the 'individuals' and their 'environment'.

The IBMS (Weed Seed Wizard, PERTH, CCCEPAE and PPUNCC) also vary widely in which details of the environment is represented. In PERTH, herbicide application is an environmental influence of plants with an effect (mortality) that is determined mechanistically and individually, depending on the genetics of the individual plant, while in the Wizard this is represented in a simpler and more empirical way. In the Wizard, the detailed weather inputs and soil condition affect the germination of seeds in a detailed and mechanistic way, while in PERTH this is represented in a simplified empirical way. CCCEPAE and PPUNCC differ from PERTH and Weed Seed Wizard in being spatially explicit, but in CCCEPAE the resolution of space is much finer, with the soil represented as 10 cm cells, while in PPUNCC the landscape is represented as 1 km cells. This means that in CCCEPAE individual plants interact with each mechanistically through short-distance dispersal and short-distance competition for soil water, while in PPUNCC, these short-distance processes must be represented more empirically, through simply capping the number of plants that can grow within a single cell. In PPUNCC the process represented in more mechanistic detail is long-distance seed dispersal.

In all of the examples presented, there were reasons why the individual-based approach was used (whether the 'individual' was a plant, or a part of a plant. One advantage of the IBM/FSPM approaches is that they are more biologically realistic, including variation (spatial, temporal, history, genotype, other) between the individuals that cannot be included in models where many intervals are pooled. This allows more aspects of reality (more relationships and interactions between components) to be included, which can then be explored in model analysis. Constructing and describing the models is often more intuitive. The representation of mechanism is deeper, in that overall behaviour emerges from the interaction of many different components. This is true even when the behaviour of the components is determined empirically, as in the seagrass and apple models. Certainly some mechanistic details can be efficiently included in models based on pooling, but as more details are included, the number of pools required increases exponentially, until it becomes more efficient and certainly easier not to pool individuals.

However, the IBM and FSPM approaches also have disadvantages. Their flexibility makes it tempting to include many different aspects of biological reality, making them more and more detailed and complex. This complexity can make them computationally slow and also difficult to analyse and understand fully, since they cannot be analysed analytically, and must be simulated for all parameter values of interest. So choices about the amount of detail to be included in a FSPM/IBM plant model should always be made with the objective of achieving the purpose of the modelling study as effectively but efficiently as possible, with only the needed amount of detail. This is why PERTH includes details of genetics and not environmental effects on germination, while the Wizard includes details of environmental effects on germination and not genetics. Similarly, this is why PPUNCC represents a large landscape at a relatively coarse scale and CCCEPAE a small area at a relatively fine scale. Or why the birch model is more mechanistic than the apple or seagrass models. In some cases, IBM models can be computationally optimised by pooling classes of individuals. This was done, for example, with the Wizard (by simplifying the soil into a number of layers and the seeds into a number of cohorts) and PERTH (for cases with a relatively small number of genes). In all cases, these kinds of design decisions should be made with a clear focus on the purpose of the modelling, rather than on simply including as many aspects of reality as possible. So the answer to our question of 'How should we model a plant?' must be 'At the minimum level of detail required to properly address the purpose of the modelling'.

ACKNOWLEDGMENTS

REFERENCES

Berger, U., Piou, C., Schiffers, K. and Grimm, V. (2008). Competition among plants: Concepts, individual-based modelling approaches, and a proposal for a future research strategy. *Perspectives in Plant Ecology, Evolution and Systematics,* 9(3-4),121-135.

Costes, E., Smith, C., Renton, M., Guédon, Y., Prusinkiewicz, P. and Godin C. (2008). MAppleT: simulation of apple tree development using mixed stochastic and biomechanical models. *Functional Plant Biology*, 35(10), 936-950.

Fourcaud, T., Zhang, X., Stokes, A., Lambers, H. and Korner, C. (2008). Plant Growth Modelling and Applications: The Increasing Importance of Plant Architecture in Growth Models. *Annals of Botany*, 101(8),1053-1063.

Godin, C. and Sinoquet, H. (2005). Functional-structural plant modelling. *New Phytologist*, 166(3), 705-708.

Grimm, V. and Railsback, S.F. (2004). *Individual-based Modeling and Ecology*. Princeton University Press.

Hanan, J. and Prusinkiewicz, P. (2008). Foreword: Studying plants with functional-structural models. *Functional Plant Biology*, 35(9-10), i-iii.

Judson, O.P. (1994). The rise of the individual-based model in ecology. *Trends in Ecology and Evolution*, 9(1), 9-14.

Renton, M., Hanan, J. and Kaitaniemi, P. (2003). The inside story: including physiology in structural plant models. *Proceedings of the 1st international conference on computer graphics and interactive techniques in Australasia and South East Asia*. New York, http://portal.acm.org/citation.cfm?doid=604471.604491

Renton, M. (2004). *Function, form and frangipanis: Modelling the patterns of plant growth* PhD Thesis, School of Physical Sciences, The University of Queensland.

Renton, M., Kaitaniemi, P. and Hanan, J. (2005). Functional–structural plant modelling using a combination of architectural analysis, L-systems and a canonical model of function. *Ecological Modelling*, 184(2-4), 277-298.

Renton, M., Guédon, Y., Godin, C. and Costes, E. (2006). Similarities and gradients in growth unit branching patterns during ontogeny in "Fuji" apple trees: a stochastic approach. *Journal of Experimental Botany*, 57(12), 3131 -3143.

Renton, M., Diggle, A. and Peltzer, S. (2008). Understanding, predicting and managing weed seedbanks in agricultural systems with the Weed Seed Wizard. *Proceedings of the 16th Australian Weeds Conference*, Cairns, Australia, 77-79. http://www.caws.org.au/awc/2008/awc200810771.pdf

Renton, M., Airey, M., Cambridge, M.L., Kendrick, G.A. (2011a). Modelling seagrass growth and development to evaluate transplanting strategies for restoration. *Annals of Botany*, accepted March 2011

Renton, M., Diggle, A., Manalil, S. and Powles, S. (2011b). Does cutting herbicide rates threaten the sustainability of weed management in cropping systems? Journal of Theoretical Biology, 283(1), 14-27.

Renton, M., Shackelford, N., Standish, R. (2011c). *Proceedings of MODSIM2011*.

Vos, J., Marcelis, L.F.M., Struik, P.C., Evers, J.B. and de Visser, P.H.B. (2007). *Functional-Structural Plant Modelling in Crop Production*, Springer, Wageningen University.