Modelling Cotton Plant Development with L-systems: A template model for incorporating physiology

[1]Hanan, J., [2]D. Thornby and [3]S. Adkins

[1]ARC Centre of Excellence for Integrative Legume Research, ARC Centre for Complex Systems, Advanced Computational Modelling Centre, University of Queensland; [2]Queensland Department of Primary Industries and Fisheries; [3]School of Land and Food Sciences, University of Queensland
E-Mail: jim@maths.uq.edu.au

Keywords: *Cotton; Gossypium hirsutum; L-systems; morphological model; carbohydrate physiology.*

EXTENDED ABSTRACT

Cotton (*Gossypium hirsutum* L.) has a complex architecture resulting from an intricate pattern of development, which strongly influences its ability to capture resources. Computational modelling can play a part in increasing our understanding of the processes intrinsic to the cotton cropping system at the level of individual branches, leaves and bolls. Such studies may be made for many different purposes, and this has a significant impact on the level of abstraction in the models of underlying function. Depending on whether the model is to be used for visualisation, for understanding the effects of processes at a population level, or for detailed examination of how physiology drives plant development, an appropriate model may be constructed with no functional interaction, with interaction between component types (*eg*. shoots and roots), or with interactions between all component modules in a plant.

L-systems (Lindenmayer 1968), a string rewriting formalism that captures growth and development as a set of rules, are the foundation of a plant modelling language called cpfg (Prusinkiewicz *et al.* 2000), which is well suited for modelling for all these purposes. This paper describes a template L-system model of the development of cotton structure that can be used to capture plant function at these levels.

Modules represent the basic components of the cotton plant: internodes I, cotyledons C, leaves L, axillary buds B, reproductive structures R, and apical meristems of the main stem A, of the monopodial vegetative branches M, and of the sympodial reproductive branches S. Each component has at least the two parameters age and node, additional parameters being added as

required. Potential inputs from models of plant function are represented in the template by named functions. Depending on the level of modelling, these may be constants, simple functions, or may also incorporate global parameters such as planting density, temperature, or resource availability.

To demonstrate the template's operation, its use in investigating the effects of defoliation (perturbing carbohydrate production) on the propensity of cotton plants to develop and extend monopodial and sympodial branches is described. In order to add physiological function to the template model, we replaced empirical rules for the growth of monopodial and sympodial branches with a system of rules relying on local pools of carbohydrate, where the values of local pools relied on acquisition in leaves and main stem translocation. Comparison of the model's performance with real plants showed that while sympodial branch extension appears to rely strongly on locally available resources, development of monopodial branches is much less so affected, in a similar way to the main stem. We also found that little or no reallocation of resources from undamaged physiological units appears to occur following defoliation in cotton.

We reached two primary conclusions from this work. First, the template model is an effective empirical framework for the physiologically explicit model of carbohydrate acquisition and branch development, and therefore is likely to be suitable for other physiological applications as well. Using the template meant not having to 'reinvent the wheel' with respect to the aspects of the cotton plant that were not of interest in the defoliation study, such as monopodial branch development.

Second, the fact that the completed model is a qualitatively accurate model of the behaviour of sympodial branches following defoliation leads us to believe that there is a functional difference between the developmental program for the two

types of branches in cotton plants. That is, sympodial branches rely on local carbohydrate resources for continued meristem activation, and

monopodial branches either do not, or do so with a much lower critical level of resource.

1. INTRODUCTION

Cotton (*Gossypium hirsutum* L.) displays a complicated pattern of morphogenesis, resulting in a complex architectural structure. Leaves vary in size and shape according to their position on the main stem and branches; branches have vegetative and reproductive forms; reproductive structures (squares, flowers, and bolls) frequently abscise; and resource demand from reproductive structures appears to result in the cessation of growth in the main stem (the phenomenon of 'cutout'). Because of the intricacies of cotton's architecture and that architecture's influence on physiological development, computational modelling can assist in increasing our understanding of underlying functional and structural processes at the level of individual branches, leaves, and bolls.

Studies of cotton's growth and development may have many different research aims, and the aims of any given study are likely to have a significant impact on the levels of abstraction and mechanism required for modelling different aspects of the plant's physiology and morphogenesis. If the structural models are to be used in scientific visualisation applications or as a platform for further simulations, an empirical model of structure that captures physiological effects under particular environmental conditions may be sufficient. If an understanding of factors affecting physiology of individual plants interacting within stands is desired, functional-structural models of aggregate activity of different component types within the plant are appropriate, for instance treating roots, leaves and stem as separate biomass compartments. If details of the physiological mechanisms driving plant development are being investigated, then the functional level must be at the component scale, for instance, incorporating the contribution of individual leaves in controlling or enabling apical development.

L-systems (Lindenmayer 1968) are the foundation of a plant modelling language called cpfg (Prusinkiewicz *et al.*. 2000) well suited for modelling for all these purposes. This paper describes a template L-system model of the development of cotton structure that can be used to capture plant function at these levels. This extends the empirical framework of previous models (Room and Hanan 1995, Hanan and Hearn 2001) to allow capture of mechanistic aspects of

development. Further, we describe the use of the template as a base model for a study into carbohydrate acquisition, translocation and allocation in cotton's fruiting branches. This study of physiology and meristem activity in branches is an example of how a physiological system can be attached to the structural template model

1.1. L-systems

The L-system formalism (Lindenmayer 1968) models plants or parts of plants as an assembly of components, each represented by a symbol with associated parameters called a module. A string of modules captures the architecture of a plant, by positioning the components relative to their neighbours, with a branching topology imposed by a hierarchy of square brackets. Growth and development of the plant structure are captured by growth rules, or productions, that are applied in parallel to the modules in the current structure in order to produce the structure at the next time step. In cpfg, productions are expressed in the form

predecessor: condition --> successor

where the *predecessor* is a string of modules with parameters expressed as variable names that will be assigned the actual values that will appear in the string, the *condition* is a logical expression that must evaluate as true for the production to be applied, and the -->, read as "produces", delineates the start of the *successor*, a string of modules whose parameters contain expressions to be evaluated before the successor is placed in the new string. In addition, pre-condition statements may be executed once the predecessor has been matched to a module in the current string, and post-condition statements may be executed if the condition then evaluates as true, and before the successor is produced. Statement types include assignment statements with standard mathematical expressions, if-then-else conditional statements, and while loops. A special form of production rule, called a decomposition, allows iterative definition of hierarchical structures. Homomorphisms are another special form of production rule used to specify graphical interpretation for visualisation purposes.

Global parameters and arrays may also be defined, and global statement blocks can be specified to include processing statements at the start and end of a simulation, and the start and end of each step in the plant's development. For complete details of

the language syntax see the cpfg manual, which is available with the L-studio package from http://www.algorithmicbotany.org.

1.2. Cotton growth habit

The cotton main stem is indeterminate and monopodial, its apical meristem producing a sequence of internodes, individual leaves, and associated axillary buds at regular intervals until the load of fruit on the plant causes vegetative growth to slow or *cut out*. The buds develop into two types of branches, monopodial vegetative branches that duplicate the growth habit of the main stem, and sympodial reproductive branches that bear flower buds (called squares), flowers, then fruit (called bolls). The first reproductive branch appears around node 6 of the main stem, and then at most or all nodes above this, with usually two to four vegetative branches appearing below that (Oosterhuis 1990). The number and vigour of vegetative branches will depend largely on resource availability, usually limited by planting density.

Unlike the main stem or vegetative branches, the reproductive branch follows a sympodial branching pattern. The development of the branch is terminated by the transformation of its apical meristem into a flower bud, after development of one compressed and one expanded internode. The compressed internode has a prophyll rather than a true leaf. The axillary bud on the second internode continues development of the branch, if sufficient resources are available. Flower buds and young fruit are subject to shedding caused by competition for resources within the same plant. Occasionally the axillary bud of the compressed internode may develop, with such branches at the first position appearing like a second branch from the same main stem node.

2. BASIC TEMPLATE

Modules represent the basic components of the cotton plant: internodes I, cotyledons C, leaves L, axillary buds B, reproductive structures R, and apical meristems of the main stem A, of the monopodial vegetative branches M, and of the sympodial reproductive branches S. Each component has at least the two parameters age and node, additional parameters being added as required. The age of all components, along with any other parameter that evolves over time, is updated in the production rules, each component having a rule like

L(age,node) --> L(age+dT,node)

where dT is the amount of time passing in the step. If the model runs on plant physiological time a convenient unit for age is the plastochron, the period between the initiations of successive leaves by the main stem apex. If the plant is running in calendar time, with daily or weekly time steps, age may be expressed in degree-days (Hanan 1997).

Potential inputs from models of plant function are represented below by named functions starting with a capital F. Depending on the level of modelling, these may be constants, simple functions, or may also incorporate global parameters such as planting density, temperature, or resource availability. In more complex situations, these inputs may be dynamically determined in global processing statements at the start of each step.

The apical meristem A drives the development of the indeterminate main stem. Once its age has been incremented in the regular productions, the following decomposition rule may be applied.

A(age,node): age>=Fplasto(age,node)
       {n_age=age-Fplasto(age,node);}
    --> I(n_age,node)[L(n_age,node)]
      [B(n_age,node)]A(n_age,node+1)

Whenever the apex is old enough, ie its age passes the age and node-specific plastochron defined by function Fplasto, it will produce a new internode, leaf, and axillary bud (collectively called a metamer), and a continuing apex. The value of Fplasto will generally increase with node number, until the boll load causes cut out to occur and the plastochron becomes very long. The temporary variable n_age is used to reset the apical clock and gives new components an age that accounts for the exact time that the transformation occurs relative to the step length. The apex A updates its node parameter in preparation for creating the next new metamer. All other components get the current node number.

The axillary buds B may eventually grow out as branches:

B(age,node): Fbranch(age,node)==DORMANT
    --> B(age,node)

B(age,node): Fbranch(age,node)==VEGETATIVE
    --> M(0,node,0)

B(age,node): Fbranch(age,node)==FRUITING
    --> S(0,node,0)

Fbranch(age,node) returns the branch type of an axial bud at each time step. Up until the time of the first fruiting branch's appearance, this function will return a value of DORMANT causing all buds to remain dormant. The first fruiting branch apex S will typically appear at around node 6, after which monopodial vegetative branch apices M may be released below that. All buds above will be released as sympodial reproductive branch apices S. The node of origin of the branch is passed as the second parameter of the new apices, in case it is needed in functional modelling of internal processes such as competition for resources.

Monopodial and sympodial branch apical processing is similar to that of the main stem for age and node.

M(age,origin,node): age>=Fm_plasto(age,node)
        {n_age=age-Fm_plasto(age,node);}
        --> I(n_age,node)[L(n_age,node)
            [B(n_age,node)]
            M(n_age,origin,node+1)

S(age,origin,node): age>=Fs_plasto(age,node)
        {n_age=age-Fs_plasto(age,node);}
        --> I(n_age,node)[L(n_age,node)]
            [S(n_age,or,node+1)]
            R(n_age,or,node)

The monopodial branch apex M reproduces the activity of the main stem. The sympodial apex, on the other hand, terminates in a reproductive structure R and branch growth is continued by the axillary bud S. Both these structures keep track of their age, main stem node of origin (or) and position on their branch (node). Plastochrons are controlled by independent functions Fm_plasto and Fs_plasto. Note that for the reproductive branch the compressed internode and its axillary bud are not explicitly modelled, but the timing of branch development is maintained by using a plastochron that is about twice as long for the sympodial branches as for the monopodial branches and main stem.

The reproductive structure R is a special case. Its development will progress through stages of square, flower, boll and open boll. The square and young boll in particular are subject to shedding:

R(age,or,node): Fshed(age,or,node) == TRUE
        --> *

where the Fshed function will return TRUE if the reproductive structure should be shed at any time, and the symbol * causes the removal of the predecessor from the string. For empirical models this function may be the result of sampling from a

probability distribution, possibly with component categorisation based on ad hoc rules (eg. Hanan and Hearn 2001), whereas for a functional model it may be a complex process involving internal competition.

Geometry is added by homomorphism rules when the modules are interpreted to create the image. For instance, there generally is a rotation of about 135 degrees between internodes along the main stem and leaf angles start low and increase as the leaf ages. Leaf shapes may be varied by position. Size of individual components is generally specified as a growth function in the homomorphism. For instance, an internode may have a rule such as

I(age,node) --> F(Fint_growth(age,node))

where F is the module used to draw a cylinder in cpfg.

3. INCORPORATING PHYSIOLOGY

This template model provides a platform for incorporating functional models into the cotton plant structure. There are a number of schemes, of differing complexity, for incorporating function into the structural template. Empirical models of plant physiology can be incorporated by gathering data and fitting functions for each of the named "F″ functions(). These may take the form of distributions or regression equations. Geometry may be determined the same way, while growth functions may be specified based on means for maximum size and duration of growth according to component categories, for example, leaf size classes determined by position. Figure 1 shows such a model based on greenhouse measurements (Room and Hanan 1995).
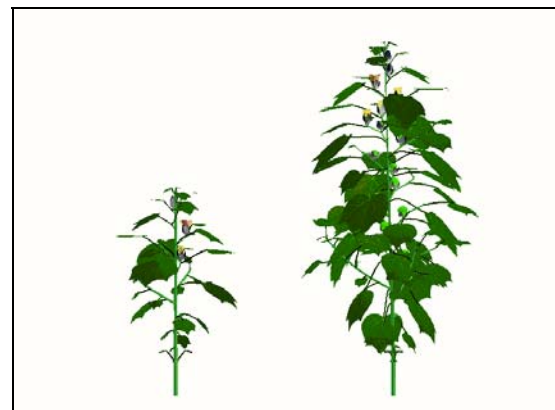
Figure 1: Two stages in the development of a virtual cotton model

Some functional models may require transport of signals through the branching structure. This can be handled using standard context-sensitive rules (Lindenmayer 1968, Prusinkiewicz *et al.* 2000). Additional parameters incorporated in internodes could carry presence/absence or continuous levels or concentrations of signal. Productions for initiation and component sensing of signals may also be required. Where global knowledge is required in more abstract models, it is often convenient to use global variables and arrays, indexed by node number, to capture information (Hanan and Hearn 2001, Thornby *et al.* 2003).

Models of physiology at intermediate levels can be incorporated using canonical models (Renton *et al.* 2003). Physiology is modelled as a number of compartments, usually component based, with fluxes between them captured in non-linear power law expressions. Constants are estimated from data. Changes in compartment levels may then be allocated to maturing components, on a demand basis, for instance.

When focusing on more complex physiological processes, some aspects may require extra detail. In models of response to damage (Thornby *et al.* 2003), this extends to the inclusion of primordia when the timing of responses depends on their presence. In this case the apices A and M produce an abstract primordia P, which has its own set of production rules to capture physiological effects. At the appropriate time, the primordia produce a metamer as described above for the apex A.

4. EXAMPLE: BRANCH PHYSIOLOGY

As an example of the template's use, a physiologically explicit approach (Thornby *et al.* 2003; Thornby and Hanan 2004) can be applied to investigate the effects of defoliation-mediated reduction in photosynthate production on the propensity of cotton to produce and extend vegetative and fruiting branches. In order to do this, a number of the above mentioned physiological functions were represented in physiologically explicit ways, such that a system of photosynthate physiology was built up and used to drive various aspects of plant morphogenesis, including branch development and extension. We were interested in the effects of local defoliation (ie. removing branch leaves and/or subtending or nearby main stem leaves) on the growth behaviours of the affected branches, particularly as mediated by changes in source-sink relationships for carbohydrate.

Photosynthate is produced in leaves, and assimilated in developing and expanding leaves and internodes, and in reproductive structures. The model makes some simplifications: only net output of carbohydrate is counted for each leaf (that is, the leaf's own respiration and maintenance requirements are assumed, not modelled) and similarly the photosynthetic activity of green squares and bolls is implied rather than modelled directly. Given that leaf photosynthesis is by far the greater source of carbohydrate to developing flowers and bolls, this simplification should have little impact on the model's validity.

Local availability of photosynthate determines the ability of leaves and internodes to develop to their full potential size, and may also result in abscission of developing flowers or bolls. Within sympodial branches, to reflect likely priorities in carbohydrate partitioning in real plants, developing (i.e. not mature) components acquire carbohydrate from the local pool (represented by the appropriate array value) in the following order for each branch: bolls, squares, flowers, leaves, internodes.

Similarly, the continuing development of sympodial fruiting branches is modelled as being dependent on local carbohydrate supplies, since this was found to be likely to be a factor in real defoliated cotton plants. As monopodial branch development was not found to be perturbed by defoliation treatments in real cotton plants, we retained the template model's obligate rules for monopodial branch extension (Thornby and Hanan 2004).

In order to test branch responses to defoliation in the cotton model, a system of carbohydrate sources and sinks, with translocation through the main stem and the branches, was set up. This photosynthate system was implemented as a two-dimensional data array initialised at the beginning of the model run (in the 'Start' statements; Prusinkiewicz *et al.* 2000), which was by definition globally available to all productions in the L-system. Using an array minimized the need for changes to the structural part of the template model. The main stem sources (mature leaves) and sinks (expanding leaves and internodes) contribute to and extract from values in the array indexed by the component's position on the main stem and by type of physiological unit – 0 for the main stem; 1 for a branch at that main stem position (and theoretically, 2 and above for additional branches at that main stem position, though these rarely appear in cotton and are not simulated in this model). Leaves contribute resources when their age is above a maturity threshold. The relevant parts of the L-system production for a main stem leaf are:

L(node, age, size, order): order==0
        {agefrac=age/MAX_AGE_LEAF;
        ps=Fphotosynthesis(age, size);
        prop_brch=Fproportion_brch(node);
        prop_ms=ps-prop_brch;}
        --> L(node,age+dd,size,axis)

That is, for leaves where order (the order of the branch or axis to which a leaf is attached) equals 0 (main stem leaves), calculate a photosynthesis amount (ps=Fphotosynthesis(age,size), where Fphotosynthesis uses a user-defined, fitted curve to relate carbohydrate output in mg to leaf age and area). The photosynthetic output is apportioned between the main stem's photosynthate pool and the pool for the local branch, where Fproportion_branch(node) is a function apportioning resources to the two pools by means of a constant. Arrays are used to keep track of pool values in the detailed programming. For branch leaves (that is, where order>=1), the production is similar, but the whole value of ps is apportioned to the local branch resource pool.

Carbohydrate produced in the leaves is translocated up and down the main stem of the plant. In order to simulate translocation, nodes N(num) were included in the model between each pair of internodes. Photosynthate can be passed both up and down the stem (as is the case in phloem transport), and the proportions of local photosynthate moved acropetally and basipetally depends on the position of the current node relative to the size of the plant – this allows photosynthate present in the lower part of the plant to be preferentially translocated towards the roots, as appears to be the case in real cotton plants (Ashley 1972, Constable 1981). The basic translocation production is:

N(num): (resource_pool)>0
        {propn_up=Fpropn_up(num);
        propn_down=Fpropn_down(num);
        res_up=resource_pool*propn_up;
        res_dn=resource_pool*propn_down;
        new_p=resource_pool-(res_up+res_dn)
            + Fresources_in(num);}
        --> N(num)

That is, where there are local main stem resources available, functions dependent on the node's position relative to the height of the main stem (Fpropn_up and Fpropn_down) determine proportional allocation of resources toward the apex and towards the root. The model keeps track of resource pool values for each main stem position in arrays, and new array values are substituted for the old ones at the end of each step. Fpropn_up, Fpropn_down and Fresources_in

control allocation of new resource values for the pools at num-1 and num+1 (that is, positions immediately above and below the current node) as well.

Translocation into and out of branches (other than during photosynthate production in subtending main stem leaves) is not modelled, as it appears to be extremely restricted in real cotton plants (Constable 1981).

In the model, each sympodial branch initiates and produces one metamer, since almost all nodes above the first sympodial branch in real plants do produce a branch (Oosterhuis 1990). However, after the first metamer, each time the sympodial branch's vegetative meristem is ready to produce the next metamer (as determined by the template model), the current level of photosynthate in the branch is tested against a minimum value. If the current level is greater than this threshold, the branch develops. If not, the branch fails to develop. If the meristem fails to produce a new metamer a set number of times, it is removed. These rules allow defoliated branches to reproduce the behaviour of real cotton plants (Thornby and Hanan 2004), without interrupting the central integrity of the template model. Furthermore, when rules for preferential photosynthate supply to developing bolls, flowers, leaves and internodes are included, the model reproduces realistic branch lengths, both in linear measurement and in number of nodes, following removal of either the subtending main stem leaf (MS), or the subtending main stem leaf plus the first branch leaf (MS+branch) (Table 1). That is, interaction between sink and source strength within each branch tends to cause the branch to stop producing new metamers as a consequence of increasing boll load, or significant defoliation, or both.

Table 1. Real and modelled branch lengths: linear measurement and number of nodes produced

Leaves cut	Real branch length (cm)	Model branch length (cm)	Real branch # nodes	Model branch # nodes
control	98.6	110.2	3.2	4
MS	47.6	68.3	1.9	2
MS + branch	44.8	32	1.0	1

The inclusion of explicitly modelled physiology for sympodial branch extension changes the treatment of this aspect of plant growth from a largely empirical one to a mechanistic one. This allows the user to investigate the effect of various

hypotheses regarding sympodial branch extension and therefore fruiting behaviours in cotton. The model is useful for investigating yield loss thresholds from defoliation damage; by changing the number and position of leaves removed in the plant, effects on boll number and size can be predicted (Table 2).

Reproductive behaviours are affected both through increasing boll abscission and by reducing the

branch's ability to produce more metamers, and thus more sites for fruiting structures.

Table 2. Real and modelled boll weights produced with and without removal of branch subtending main stem leaf

MS leaf	Real boll dry mass (g)	Model boll dry mass (g)
intact	98.6	110.2
removed	47.6	68.3

5. CONCLUSIONS

This L-system-based structural model, extended with basic geometric parameters, forms a solid template supporting many levels of functional modelling. This reduces the need to redo basic morphological work for every new model, while giving a framework for collection of basic structural information for new varieties**.** The researcher is freed to look at more complex issues of physiology and growth processes in general. This complexity may be reduced by modelling less important functions empirically, while building detailed mechanistic models in an area of interest, such as the interaction between defoliation and local branch growth studied here. The detailed, mechanistic model of sympodial branch growth as mediated by carbohydrate supply has allowed us to examine the problem-space of how defoliation might affect the propensity of branches to develop, extend, and support reproductive structures. Building this mechanistic model on the empirical framework of the cotton template model allowed us to investigate the effects of defoliation without needing to build and verify details of a stand-alone cotton model. This also improves our confidence in the model as a whole and reduced the time taken to build the mechanistic model significantly.

Future work on the template could include development of detailed reporting options, set up to produce standard statistics on component counts and timing of development that will verify new

functional additions have not disrupted previously modelled processes. This will become more important if integrated, detailed models of function are required. Another addition may be individual component primordia, particularly if detailed models of long distance signalling and genetic regulatory networks are to be included.

6. REFERENCES

Ashley, D.A. (1972), C-labelled photosynthate translocation and utilization in cotton plants. *Crop science* 12: 69-74.

Constable, G.A. (1981), *Carbon fixation and distribution in cotton.* PhD dissertation, Australian National University, Canberra.

Hanan, J. (1997), Virtual plants – integrating architectural and physiological models, *Environmental modelling and software*, 12(1): 35-42.

Hanan, J. S. and Hearn, A. B. (2001), Linking physiological and architectural models of cotton. *Agricultural systems* 75 (1): 47-77.

Lindenmayer, A. (1968). Mathematical models for cellular interactions in development, Parts I and II. *Journal of theoretical biology* 18, 280–315.

Oosterhuis, D.M. (1990), Growth and development of a cotton plant. *Nitrogen nutrition in cotton: practical issues*. ASA.

Prusinkiewicz, P. (1999), A look at the visual modeling of plants using L-systems, Agronomie: *Agriculture & environment* 19: 211-224.

Prusinkiewicz, P., Hanan, J.S. and Měch, R. (2000), An L-system-based plant modeling language. (ed. Nagl, M., Schurr, A. and Munch, M.) *Lecture notes in computer science 1779: Applications of graph transformation with industrial relevance*. Berlin: Springer-Verlag, 395-410.

Renton, M., Hanan, J., and Kaitaniemi, P. (2003), The inside story: including physiology in structural plant models, *Proceedings of GRAPHITE 2003*, 95-102

Room, P. M. and Hanan, J. S. (1995), Virtual cotton. In Constable, G.A. and Forrester, N.W. (eds): *Challenging the future: proceedings of the world cotton research conference 1*, CSIRO, Melbourne, 40–44.

Thornby, D., Renton, M. and Hanan, J. (2003), Using computational plant science tools to investigate morphological aspects of

compensatory growth. *Lecture notes in computer science* 2660:708-717.

Thornby, D. and Hanan, J. (2004), A modelling exploration of branch extension, defoliation responses, and carbohydrate physiology in cotton (*Gossypium hirsutum* L.), in Godin, C. (ed) *4th International workshop on functional-structural plant models*, 7-11 June 2004 Montpellier, France, 319-322