

StemGL, a FSPM tool dedicated to crop plants model calibration in the single stem case

Fabienne Ribeyre

CIRAD, Univ Montpellier, UPR Bioagresseurs, F-34398
Montpellier, France

Marc Jaeger*, Philippe de Reffye

AMAP, Univ. de Montpellier, CIRAD, CNRS,
INRA, IRD
Montpellier, France
marc.jaeger@cirad.fr

Abstract— FSPM are getting popular, open to a wide range of application and implemented in numerous simulation software or platforms. However, those seldom cover the field calibration aspects. We propose here a simple tool, StemGL, limited to single-stemmed plants, covering both calibration and simulation aspects of biomass production and allocation. Based on GreenLab model assumptions, the tool implements stochastic simulation capabilities; it offers virtual insights, and also presents field data comparisons. The application is delivered with a set of examples, with virtual plants to analyze and simulate, and real field plants with their parameters to retrieve.

Keywords—FSPM; software; calibration; simulation; GreenLab; Octave; data fitting; tomato

I. INTRODUCTION

Within the plant modelling community, the FSPM approaches developed for several decades since the middle 80', becoming mature and opening to applications [1]. The spread and the development of the FSPM models are strongly correlated to their implementations in specific software applications or software platforms. In fact, despite common shared notions, a huge variability of tools raised from the community. The way structure is represented and built, the way biomass allocation and transportation is operated, the technical environment, the genericity of the tool and most of all, the aim of the tool (simulation, calibration...) are some contributors to this diversity.

However, few tools are dedicated to or integrate parameter estimation. This is specifically true for models involving a detailed structural description such as AmapSim [2], Lignum [3], L-Peach [4], GroGra (implemented in GroImp [5]), and QingYuan [6]. And more curiously, such is also the case for platforms hosting models, such as GroImp [5], Amapstudio [7] and OpenAlea [8]. However some approaches did develop calibration tools, such is the case of GreenLab early studies with GreenScilab [9], and later with Gloups [10]. Their underlying fully mathematical expression makes the model more easily be reversed. But further due to their genericity and environment (Scilab), the spread of these tools is rather limited and confidential since quite complex and heavy. Finally, the Pygmalion platform [11] is obviously the first platform dedicated to plant model comparisons, parameter estimation

and sensibility analysis, but does not host real FSPM model yet.

In this context, we propose here a simple tool covering both calibration and simulation aspects of biomass production and allocation. Called StemGL, the application limits to single-stemmed plants, involving a limited set of parameters. Based on GreenLab model assumptions, the tool implements stochastic simulation capabilities and biomass reallocation; it offers virtual insights, and also presents field data comparisons. The paper organizes as follow:

The first section describes the models assumptions, input and fitting principals. The technical aspects follow with examples of calibration and simulation descriptions. In a second section some simulation and fitting examples are given before concluding with the ongoing developments.

II. STEMGL PRINCIPLES

The StemGL principles inherit from the Gloups software [10] and GreenLab model [12]. Applied to single stem plants, structural aspects are strongly reduced, but some specific points are more detailed.

A. Model assumptions

Considering the structural aspects, only one physiological age is used since the plant has only one stem. Moreover, the modeling of the meristem functioning process is limited to the growth because viability and branching are not considered.

Considering the functional aspects, StemGL introduces a reserve common pool considered as an organ, competing with the others. The model considers thus 8 types of organs (leaf limb, leaf petiole, common pool, internode, growth rings, female fruit, male fruit, and roots). The organ expansion duration may vary according to the organ ontogenic age (the age of the plant at organ appearance date). The model allows biomass reallocation from the petiole, the leaves, and the internodes.

As the classical GreenLab model, StemGL allows varying the environmental conditions (through an aggregated parameter in the Lambert-Beer law equation).

B. StemGL parameters

In this new tool, eight categories of parameters are considered to model plant growth. The two first concern structural aspects, the others the functional ones.

P1. The development parameters. They include the Bernoulli probability for developing, number of random repetitions, the number of growth cycles, and the final age of organogenesis.

P2. The number of occurrences per phytomer of each organ type (considering that leaf limb and petiole occurrences are the same). We define for each active organ type the following functional parameters:

P3. The functioning duration and the organ expansion duration. Expansion duration value may vary according to organ ontogenic age (the age of the plant at organ apparition). The application allows four steps of expansions durations (except for the reserve pool and the tape root).

P4. The expansion time. P5. The expansion delay.

P6. The sink. Sink functions define the demands of organs. These depend on organ age (constant or beta law). They are characterized by their strength. Sink functions are proportional each to the other and can be different for leaves and internode when wearying an inflorescence. By convention, leaves sink strength is set to 1. For leaf, petiole, and internode, 2 extra parameters can define a remobilization function.

P7. Leaves and internodes dimensions. These are defined using their biomass and dimensional allometry factors.

P8. Global functioning parameters include constants radiation and climatic efficiency, production surface, seed initial weight and parameters for seed and pool emptying.

C. StemGL fitting parameters

The user may define the list of parameters to be fitted among the sinks and expansion parameters. He will also have to fill up a target file, prefilled by the application.

This target file comprises the following information: the maximum age of the plant, ages at dates of measurement, organ cumulated weight and organ individual weight and rank along the stem. At each measurement date, the total leaf and internode weights should be reported, others organs summarized reports are optional. Detailed organ weights and ranks (i.e. positions given in growth cycles from the stem tip) have not necessarily to be specified for every phytomer, depending on the parameter to be estimated. In practice, missing organs can also be reported (setting their weight to 0), and they are differentiated from the not measured ones.

In order to report real data from fulfilled, for specific plants were all organ types are not allowed at each phytomer, a mask file is provided that includes 0 or 1 depending on the possibility to have the specified organ at each phytomer.

D. StemGL fitting procedure

For fitting, the user performs a 3-step procedure, the last one being iterative. Only biomass production is fitted, not the development parameter (the Bernoulli probability can be retrieved straightforward from the phytomer distribution mean/variance ratio [13]).

Step 1. Data collection. The collected data are usually classic agronomic traits such as weight, number, and rank along the stem of different organs. Two ways are possible to calibrate: mono and multistage data. In this latter case, parameters values can be fitted using several observed ages measurements. However, this is not a requirement, the latest observation stage is sufficient. Adjustment can be done on compartments, on organic series [14] (i.e. along the stem in our case) and on both. Data are usually collected on a plant population, at various observation stages. Follow-up is sometimes impossible since some traits are easier to collect by destructive collection (such as root weights and dimensions).

Step 2. Known and measurable parameters. Plant expert knowledge is important to determine which parameters to adjust and which parameters are known or easily measurable. Some measurable functional attributes can be set to their mean values, such as allometry parameters, expansions, and functioning duration. Allometry parameters are deduced from the collected organ dimensions. Those parameters are considered as constant and are classically estimated from mean statistics on leaves and fruits.

Step 3. Fitting the functional hidden parameters

Functional parameter fitting is performed iteratively with measured data corresponding to precise development observation stages. The hidden functional parameter retrieval procedure classically follows this workflow:

Hidden parameters are filled using default values.

The parameters to be identified are then fitted using the generalized least square method, usually in several steps starting from the production equation.

The first level estimated are usually the global functioning parameters. Organ sinks are then estimated as a single ratio of leaf sink. Beta laws are then defined for biomass partitioning. Lastly, the secondary growth is estimated (if required).

The fitting iterates on the target file, updating the hidden parameter values until the distance between the simulated traits and the measured ones stays constant.

E. StemGL simulation code

The simulation process in StemGL first computes the stochastic structural development (i.e. organ positioning). A mask may be applied to it, avoiding specific organ positions. In a second step, biomass production and allocation are simulated including, if necessary, variable sink strength, remobilization, and environmental conditions changes. Simulations are done at different growth cycles, specified by the user, corresponding to different observation stages. An observation file is generated and can be used later as a target file for fitting.

F. Technical aspects

StemGL was written in Matlab programming code, and tested under the free GNU-Octave [15] environment. The code is rather compact, composed of 50 source files, each corresponding to a specific function. The interface is so far limited to graphical outputs.

Parameter files, mask files, and target files are ASCII files, offering files Excel compatibility for clear data exchanges.

The application is freely available from the GreenLab website <http://greenlab.cirad.fr/StemGL/>. The package offers the simulation code, a parameter index, and a set of example files. A tutorial is under construction.

G. StemGL examples files

In the next Result section, we apply the fitting on real tomato field data. The final parameter file is built using an initial one, and a mask.

Tomato crop field measurements [16] include the global measure of biomass at 9, 17 and 28 cycles, for 5 organ types: Blade (leaf limb) and petiole, internode, female fruit, and root. Some parse detailed information of organ biomass per rank (along with the stem) is given for leaf limbs, petioles, internodes, and female fruits.

The mask file is fulfilled allowing the occurrence of all organs at every rank except for female fruits allowed only at the observed ranks.

The number of iterations in the calibration process is fixed within the tool (classical around 10).

We first calibrate the parameters defining a reference set. This set is then compared to another one, in which we reduced the fruit expansion time (from 30 to 10 cycles) (set2), and a last one, in which we reduced the fruit sink strength (from strength 25 to strength 15) (set3).

Since this example on tomato is a not stochastic case, we also illustrate the stochastic capabilities of the application on a random stem.

III. RESULTS APPLIED ON FILE EXAMPLES

The following screenshots illustrate StemGL runs on the Tomato data fitting (the tool was also used on maize and beetroot, sunflower field data).

Typical result output some first some global results at the plant level such as the plant demand, its production, and the number of organs per cycle. The overall quality of the simulation can be grabbed with a synthetic plot showing the cumulated distance between the estimations and the measured observations (Fig 1, right).

By default, the application plots the cumulated biomass allocated to each organ type at each growth cycle; and for each observation date, the production per cycle (Fig.1, right).

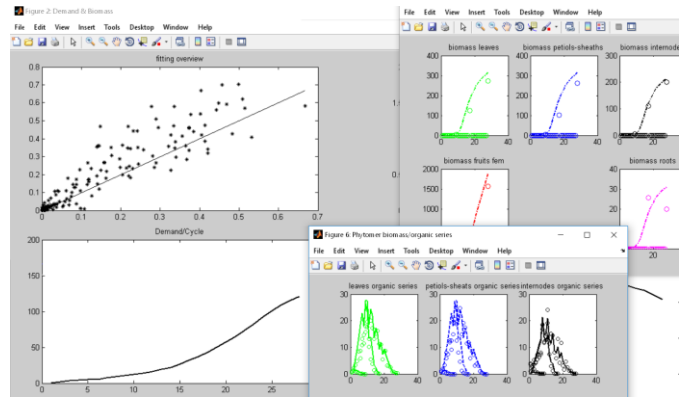


Fig. 1. StemGL output view after calibration. Left top: the overall biomass estimation error is plotted. Left bottom: global plant demand according to the growth cycle. Right Top: organ compartments cumulated biomass. Right bottom: biomass allocated to organs according to the growth cycles for the three observation dates. The simulated values are plotted as curves, while the observed values are circles. Green color stands for leaves, blue for petiole, black for internodes, red for fruits, and pink for root.

Fitted parameters and input data can also be plotted. Such are the cases of the environmental condition factor, the expansion and the sink functions. The simulated plant corresponding to the target adjustment is also generated and displayed.

The figures (Fig.2 to Fig.6) illustrate the curves generated with the reference tomato data (set1), the reduced fruit expansion one (set2) and the reduced fruit sink strength one (set3)

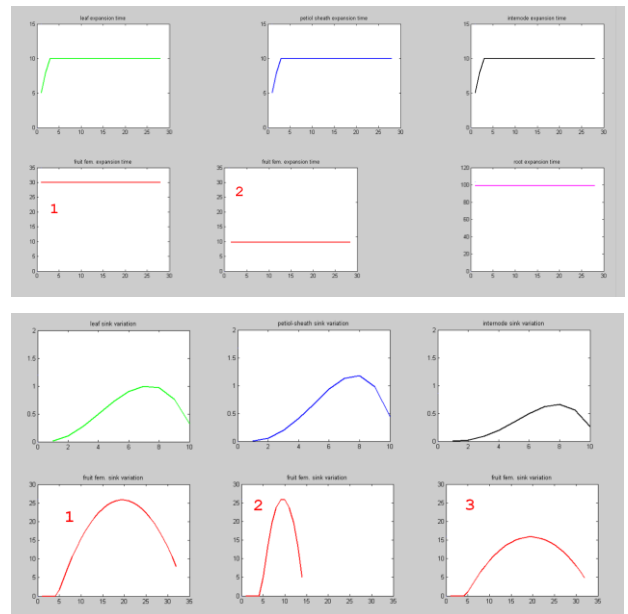


Fig. 2. Visualizing parameters in StemGL. Top: Expansion times on the tomato data sets. These are the same for all sets and organs, except for fruits in the set2, where the expansion duration was dropped to a third. Bottom: Sink strength functions on the tomato data sets. These are also the same for all organs and sets, except that the fruit sink strength was nearly divided by 2 in the third set. Note on set2 that the sink strength shape squeezes due to the expansion reduction.

The resulting plant structures are presented below (Fig.3).

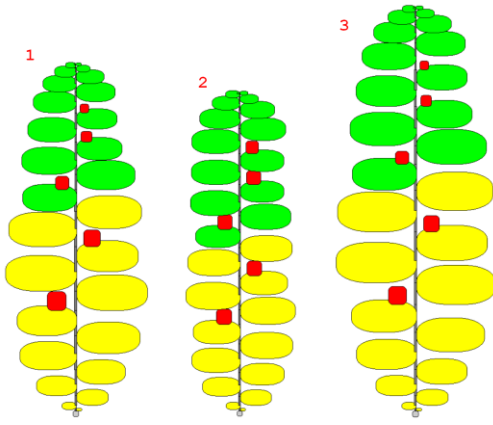


Fig. 3. The three tomato sets plant representations generated by StemGL. Note the light organ dimension changes. Functional leaves are colored green while the others (older) are yellow.

At the plant level, the biomass demand and production exhibit light change along time (Fig.4).

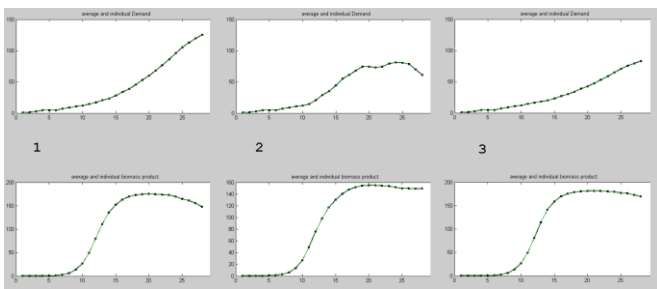


Fig. 4. Global Plant comparisons on the three tomato sets. The upper row shows the plant demand, the lower the cumulated biomass. Demand increases over the growth in reference (set 1). Form remain similar in set3 but at a lower level. In set2, the demand gets stable and then decreases since the fruit demand, fixed at a high-level, drops. Cumulated production curves are similar both in shape and values.

At the plant level, the result interpretation is not obvious, insights into functioning are necessary. We focus here on leaves and fruits.

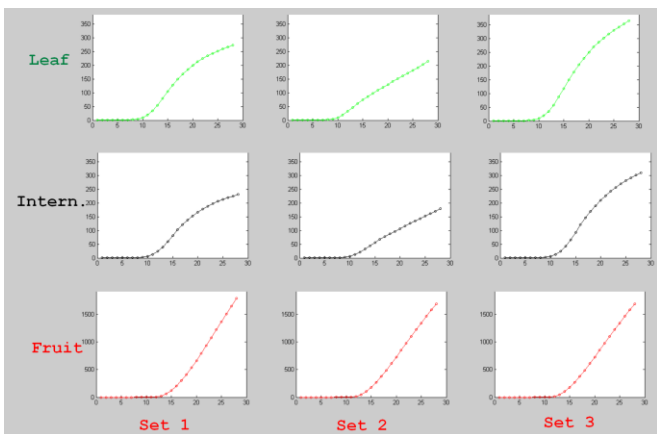


Fig. 5. Dynamics in the organ compartments. Dynamic of biomass production allocated to leaves and fruits show differences between the sets (slightly fewer fruits and leaves in set2; more leaves and fewer fruits in set3). Internodes compartment dynamics is similar to the leaf (and petiole) ones.

Organ compartment output corresponds to a classical production monitoring. This level is however still not the best one for the biomass compartment understanding. As shown below (Fig. 6), the organic series is more informative.

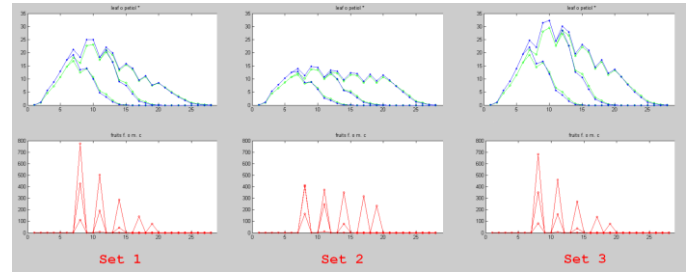


Fig. 6. Leaf and fruit organic series on the three tomato sets at observation dates 9, 17 and 28. The upper row shows the limb (and petiole) biomass production, the lower the fruit ones. Biomass variations are linked to the fruit expansion duration limitation in set2 and the sink strength decrease in set3. On all sets, the biomass of functioning leaves increases till 8 cycles which is the appearance of first fruit. After cycle 10, functioning expansion of leaves, at each cycle, aged leaves stopped their expansion but continue to produce biomass 5 cycles more as leaves functioning duration is 15 cycles. In set 1 and set 3, the increase in the number of leaves does not compensate for the stop in aged leaves production surface and leaves biomass decreases after cycle 10. In set 3, leaves are bigger than in set 1 because fruits compete less for biomass. In set 2, production surface is almost stable after cycle 10 because each fruit consumes during less cycles so overall fruit consumption is lower.

The last figure (Fig. 7) illustrates some of the stochastic capabilities offered by StemGL. On this example, the development follows a binomial law with a probability b to occur set here to value 0.8. The stochastic mode is automatically generated when b is lower than 1 and when the number N of stochastic individuals is greater than 1 (N is set to 10 in this example).

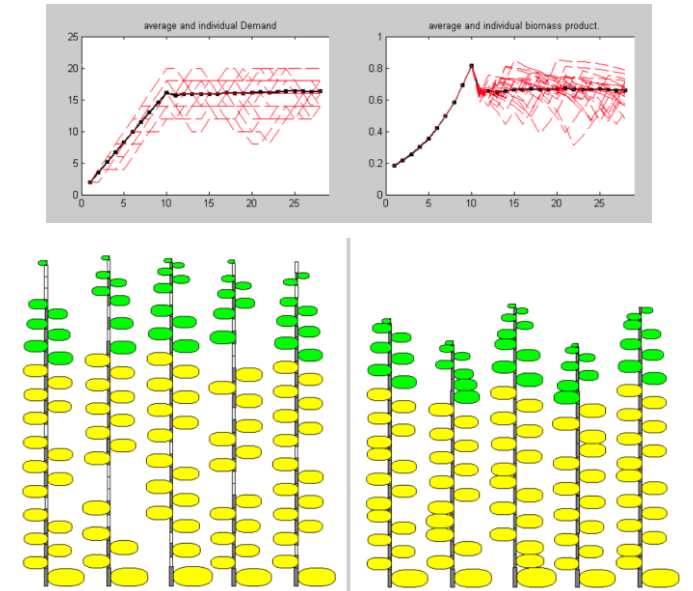


Fig. 7. Stochastic representations in StemGL. Top: plants individual (red) and average (black) demands and biomass. Bottom left: 5 constructions among the 10 simulated represented with the pauses (empty phytomers). Right, their respective topographic representations, corresponding to observations

At global levels, unlike organic series levels, the N curves are displayed with the average one. The user may also select the number of stochastic individuals Np among the N ones to be represented (Np is set to 5 here). Representations may include (or not) the breaks (Fig. 7).

IV. CONCLUSION

We presented here StemGL, a FSPM tool fitting and simulation tool dedicated to single stem plants, based on the GreenLab formalism.

The application can generate single and population individuals (stochastic case). The output generated are fitting results, fitting targets and masks at various observation dates, biomass quantity, and simulation output graphs including structure representation with and without breaks. It offers new insights of fitted functions and cumulated agronomical traits. This tool is available for the community, running under Matlab and GNU Octave environments.

The tool operated with a simple dataset frame of 8 parameter item sets, whatever the species considered. Datasets with adjusted parameters are already available for various agronomic plants (beetroot, maize, sunflower, tomato...).

Several threads of future works are in progress and further developments are planned: development of an input interface and integration under other environments, R in particular. We aim to facilitate the tool appropriation by agronomists and facilitate its evolution.

REFERENCES

- [1] Y. Guo, T. Fourcaud, M. Jaeger, X. Zhang, B. Li. "Plant growth and architectural modelling and its applications". In *Annals of botany*, 107(5), 2011, pp. 723-727.
- [2] J.F. Barczy, H. Rey, Y. Caraglio, P. de Reffye, D. Barthélémy, Q. Dong, T. Fourcaud. "AmapSim: a structural whole-plant simulator based on botanical knowledge and designed to host external functional models". In *Annals of Botany*, 101(8), 2007, pp. 1125-1138.
- [3] J. Perttunen, R. Sievänen, E. Nikinmaa. "LIGNUM: a model combining the structure and the functioning of trees." *Ecological modelling* 108.1-3, 1998, pp. 189-198.
- [4] M.T. Allen, P. Prusinkiewicz, T.M. DeJong. "Using L-systems for modeling source-sink interactions, architecture and physiology of growing trees: the L-PEACH model". In *New Phytologist*, 166 (3), 2005, pp. 869-880.
- [5] O. Kniemeyer, G. Buck-Sorlin, W. Kurth. "GroIMP as a platform for functional-structural modelling of plants". In *Frontis*, 2007, pp. 43-52.
- [6] J. Hua, M.G. Kang, P. de Reffye. "An Interactive plant pruning system based on GreenLab model: Implementation and case study". In *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on. IEEE. Vol. 4, 2001, pp. 185-188.*
- [7] S. Griffon, F. de Coligny. "AMAPstudio: an editing and simulation software suite for plants architecture modelling". In *Ecological Modelling*, 290, 2014, pp. 3-10.
- [8] C. Pradal, S. Dufour-Kowalski, F. Boudon, C. Fournier, C. Godin. "OpenAlea: a visual programming and component-based software platform for plant modelling". In *Functional plant biology*, 35(10), 2008, pp. 751-760.
- [9] M.G. Kang, X. Wang, R. Qi, P. de Reffye. "GreenScilab-Crop, an open source software for plant simulation and parameter estimation". In *Open-source Software for Scientific Computation (OSSC), 2009 IEEE International Workshop on IEEE, 2009, pp. 91-95.*
- [10] P. de Reffye, M. Jaeger. "Modèles mathématiques du développement et de la croissance de l'architecture des plantes. Le cas du modèle GreenLab". In Varenne, F., Silberstein, M. (Eds) *Modéliser et simuler. Epistémologies et pratiques de la modélisation et de la simulation. Tome 1. Paris : Editions Matériologiques(Sciences et Philosophie, vol. Sciences et Philosophie). 2013, pp. 625-658.*
- [11] P.H. Courmède, Y. Chen, Q. Wu, C. Baey, B. Bayol. "Development and evaluation of plant growth models: Methodology and implementation in the Pygmalion platform". In *Mathematical Modelling of Natural Phenomena*, 8(4), 2013, pp. 112-130.
- [12] H.P. Yan, M.G. Kang, P. De Reffye, and M. Dingkuhn, "Adynamic, architectural plant model simulating resource-dependent growth," *Annals of Botany*, vol. 93, 2004, pp. 591–602.
- [13] P. de Reffye, M. Cognée, M. Jaeger, B. Traore. "Modélisation stochastique de la croissance et de l'architecture du cotonnier". 1. Tiges principales et branches fructifères primaires. In *Coton et Fibres Tropicales*, 43 (4), 1988, pp. 269-291
- [14] M.Z. Kang, J. Hua, P. de Reffye, M. Jaeger. "Parameter identification of plant growth models with stochastic development". In : 2016 IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA). IEEE, FSPMA. Qingdao : IEEE, 2016, pp. 98-105.
- [15] GNU Octave, 2018. <https://www.gnu.org/software/octave/>
- [16] M.G. Kang, L. Yang, B.G. Zhang, P. de Reffye. "Correlation between dynamic tomato fruit-set and source-sink ratio: a common relationship for different plant densities and seasons?" *Ann Bot.* 107(5), 2011, pp. 805-15.