Activity Report 2011

Team DIGIPLANTE

Mathematical Modelling of Plant Growth

Table of contents

Team DIGIPLANTE

Keywords: Plant Development, Plant Growth, Statistical Learning, Stochastic Modeling, Optimization

Digiplante is a joint team INRIA Saclay - Ecole Centrale Paris, located at the MAS laboratory, Ecole Centrale Paris.

1. Members

Research Scientist
Samis Trevezas

Faculty Members
Paul-Henry Cournède [Team Leader, Professor, Centrale Paris, HdR]
Véronique Letort [Associate Professor, Centrale Paris]

PhD Students
XiuJuan Wang [Centrale Paris, CAU]
Cédric Loi [Centrale Paris]
Qiongli Wu
Zhongping Li [Centrale Paris]
Fenni Kang [Centrale Paris]
Charlotte Baey [Centrale Paris]
Yuting Chen [Centrale Paris]

Administrative Assistant
Christine Biard

Others
Benoît Bayol
Claire Mouton [Centrale Paris]
Chi-Thanh Nguyen [Centrale Paris]
Corina Iovan

2. Overall Objectives

2.1. Research fields

Plants play a major role in our society for human food, biodiversity, as environmental regulator, materials or energy source. The human management of plants is therefore a crucial issue for the future of our societies, and the way this management is conducted today is far from satisfactory: pollution by intensive farming, agricultural shortage in some parts of the world, environmental catastrophe caused by extreme deforestation, loss of biodiversity. And yet, plants are organisms that are well known by biologists. This important knowledge has only poorly allowed the development of management tools, probably from the lack of a systemic approach. Mostly empirism and Fisher's statistics have been used to help define farming, forestry or breeding strategies.

The first models of plant biophysics only appeared in the seventies (Monteith, 72,77; De Wit, 78). At this stage, energetic budgets are considered at field scale. In parallel, mostly for computer graphics applications, individual-based models of plant geometry were developed, based on formal grammars (Smith, 84; Prusinkiewicz, 88,89) or automata (De Reffye et al., 88). The need for applied mathematics remained quite limited at this stage (numerical simulation, computer graphics). The biophysical and geometrical approaches would finally meet in the late 90's and see the functional-structural plant models (FSPM) community emerge: interactions between plant structure and biophysical functioning is henceforth seen as the key to understand plant growth and offer a framework for the integration of existing biological knowledge. The mathematical issues raised by this new type of models are spectacular. Plants are now described as full systems, both structurally and functionally, and appear as typical complex systems: multi-scale issues, multiple and interacting biophysical processes, regulation and retroaction loops, that can no longer be fully studied with only simulations.

The Digiplante team was born in this context around Philippe de Reffye in 2002, first incubated in the Metalau group at INRIA Rocquancourt (M. Goursat, J.P. Quadrat) and as an independent joint team between INRIA, Centrale and CIRAD in 2004, with P. de Reffye and P.H. Cournède as permanent researchers. The main objective of the team at this time was the development of the GreenLab model [De Reffye and Hu, 2003], its mathematical formalism and the fitting of real plant data to study the genericity and applicability of the model. New concepts are introduced, like the interactions between structure and functioning, the effects of competition for light or a new model of secondary growth. The Digiplante software developed by Centrale propose a versatile fitting tool, that is used to study a wide variety of real plants (sugar beet, maize, rice, sunflower, Arabidopsis, rapeseed, cucumber, sweet pepper, tomato, cotton, grapevine, beech tree, pine tree, cecropia) with different biological partners. In 2007, Marc Jaeger joined Digiplante after 5 years of LIAMA's direction. He brought with him the computer graphics and plant visualization themes in Digiplante. In 2007, the team was also transferred to INRIA Saclay, it became an 'Equipe-Projet' INRIA in 2008 and was evaluated positively in 2009.

When confronting the GreenLab model to real plant data, the chosen approach was mostly horizontal, that is to say that the general idea was to see if the concepts implemented in the GreenLab model could apply to a wide variety of plants. This strategy allowed the implementation of new features in the model, making it suitable for plants as diverse as Arabidopsis and Beech tree. This step can be seen as successful. However, in only very rare cases the robustness of the model and its predictive capacity have been studied. When implemented, this vertical approach was not fully satisfactory, which makes the model still unproper for real applications: modeling the interactions with the environment should be improved, no measure and control of model uncertainty is provided. The reason for this is still the lack of a proper mathematical methodology for the analysis of models and their full evaluation.

Digiplante is now a joint team between INRIA Saclay and Centrale, under the scientific responsibility of P.-H. Cournède. The objective of Digiplante is now clearly orientated towards the development of a proper mathematical methodology for the analysis of models and their full evaluation, around 3 axes:

- Model Design and Analysis
- Model Evaluation from Experimental Data and Model Selection
- Methods for the Applications: Prevision, Control and Genetic Improvement

2.2. Highlights

- PhD defense of Wang Xiujuan, April 2011
- PhD defense of Cédric Loi, May 2011

3. Scientific Foundations

3.1. Model Design and Analysis

3.1.1. From biological concepts to equations

This step corresponds to the writing of the model equations based on the agronomic / biological knowledge. It will be done in close collaboration with the partner institutions. We will continue working on the GreenLab model, and will also consider other family of models (STICS INRA-Avignon, NEMA INRA-Grignon, SUNFLO-CORNFLO-SOYFLO Syngenta ...). 3 specific points are now our priorities:

- better integration of the environment (specifically water and Nitrogen). This is still poorly taken into account in the GreenLab model, and is quite restrictive for model applications.

- modeling plant populations from the individual-based model, by studying competition between plants and the inter-individual variability,

- modeling the genetic determinism of parameters. In a perfect model, we would write $\frac{dX}{dt} = F\left(X(t), P, E(t)\right)$ where $X(t)$ are the state variables (masses of plant components), $E(t)$ represent the environmental factors (radiation, temperature, soil water content ...) and P are variety-specific parameters from genetic origins. What would be very interesting is to write $P = H(G)$ where G represents plant genetics. Several methods are possible, including metabolic networks, but in a first step we will consider methods derived from quantitative genetics.

3.1.2. Mathematical formalism

- Formal grammars and combinatorics: in the last two years, in the PhD thesis of Cedric Loi [Loi et al. 08,10,11] some very precious results have been obtained, linking the former formalism in GreenLab (dual-scale automaton) to the theory of formal grammars (L-System). In the stochastic case, the link with branching processes was also studied, which allowed the computation of moments and generating functions for the numbers of elements in plants. In collaboration with J. Françon (Univ. Strasbourg), symbolic methods derived from the combinatorics approach of Flajolet also allowed the computation of the generating functions of the occurences of patterns in plants. Such results led to the definition of new methods to estimate the parameters of stochastic models of plant organogenesis. This promising approach still needs to be explored: extend the cases in which the distribution of patterns can be derived, comprehensive study of the estimation methods.

- Continuous models of plant growth, time-delay systems: traditional models of plant architectural growth (like GreenLab) adopt a discrete formalism (based on the discrete steps in the automaton or grammar theory defining architectural growth cycles). It proves limiting when considering plant-environment interactions. Therefore, a continuous version of the GreenLab model has been derived [Li et al., 2009], at least for the functional parts. It raises interesting numerical issues (discretization schemes and optimal control for time-delay systems). Moreover, the structural part is not yet written in a continuous way. Current studies are carried out.

3.1.3. Mathematical and statistical analysis of model structures

When model equations are written, a fundamental step is their mathematical analysis: limit and stability analysis, identifiability, sensitivity and uncertainty analysis. A few important results have been produced by Digiplante on this aspects (conditions for the generation of rhythms [Mathieu et al., 2008], designing specific methodology for the global sensitivity analysis of functional-structural plant models [Wu et al., 2011]). One of the key points to explore concerns the study of complex systems: plant integrative modeling (especially functional-structural plant modeling) implies different scales of biophysical processes, some of them are particularly well-known, but rarely the interactions between these processes when considering more global phenomena at plant or field scale. Global sensitivity analysis offers very interesting perspectives to study such integrative models (as well as some linked methodologies: model reduction / meta-modeling). A collaboration with one of the major group in the world (Saltelli, Tarantola in the Joint Research Centre (JRC) of the European Commission, Ispra Italy) is starting about the results of Qiongli Wu's PhD.

3.2. Model Evaluation from Experimental Data

This point can be considered as central in the DIGIPLANTE project.

3.2.1. Parametric Identification : estimation of model parameters and evaluation of estimation uncertainty

- Theory and methods. The parametric estimation of the GreenLab model was classically done with estimators of the Generalized Least Squares type. It assumed a diagonal covariance matrix for the errors of the model outputs. If it has provided efficient algorithms to determine parameters ensuring a satisfying goodness-of-fit, it proved restrictive to study model robustness and determine confidence intervals for the parameters, which is crucial for applications. A new perspective is currently studied to improve estimation and uncertainty characterization. An equivalent description of the dynamic system in the framework of hidden (latent variable) models was formulated (cf. CournÃ¨de et al., 2011). Statistical estimation in this framework can be tackled with tools borrowed from the theory of hidden Markov models, such as maximum likelihood estimation. In our case, the associated likelihood function cannot be computed in a closed form. Simulation based methods are in progress in order to implement proper stochastic versions of the EM algorithm and stochastic gradient methods for state and parameter estimation. In this direction, the class of sequential Monte-Carlo, particle filter and MCMC algorithms, can be used for maximum likelihood estimation and seems particularly adapted to our case. A collaboration with the Univ. of Patras is starting on this issue. The same type of methods can be used in Bayesian inference. It is also explored, for situations in which priors are easy to determine (study of genetic populations, data assimilation ...).

- Application to real plants: this aspect has been one of the strong points of Digiplante: a wide variety of plants have been studied with the GreenLab model, always confronting the model to experimental data. This study will continue, with the double objectives of improving / validating / comparing models and testing our estimation methods. However, it is important to focus on the plants for which we have rich data sets, allowing a proper model validation (with training and testing data sets). The collaborations with ITB (for sugar beet), INRA-Grignon (for rapeseed), Supagro Montpellier (for Sunflower and Grapevine) Cirad-Guyana (for Cecropia) and China Academy of Forestry (for pine) are long-term partnerships that make it possible to get these good data sets on different types of plants, with different levels of difficulty.

3.2.2. Model selection

In plant growth modeling, it seems that each research group is developing its own model. It is thus crucial to compare, conceptually and mathematically, the existing models, in order to assess their differences and select the 'best' models regarding specific objectives. Therefore, several classical models (STICS, PILOTE, ADEL-NEMA, SUNFLO / CORNFLO ...) are also considered in IPANEMA beside the GreenLab model. Our objective is to test different selection criteria, particularly MDL (Minimum Description Length) in collaboration with L2S Supélec-CNRS and MSEP (mean-square error of prediction).

3.2.3. Optimization of experimental protocol for phenotyping

If we obtain a good estimation of the uncertainty in model parameters (that is the objective of the research axis described in 3.2.1), we will also be able to optimize the experimental protocols. This is particularly important in phenotyping for seed companies, that need to evaluate the performances of large numbers of new varieties each year. The optimization concerns the amount of data to collect in a given experimental situation, and the number of experimental situations (with respect to climatic scenarios). The PhD of Fenni Kang studies these questions, in collaboration with J. Lecoeur (Syngenta).

3.2.4. Data acquisition from aerial images and data assimilation

Using real data is the key to decrease model uncertainty. For this purpose, aerial (or satellite, or drone) images provide a very interesting source of information. A new researcher (Ingénieur Confirmé) in the group Corina Iovan is a specialist of image analysis for vegetation. The objective is to assimilate this data, in order to:

- characterize plant population (species, positions, functional characteristics ...),
- correct / improve model prediction. Links with works in data assimilation developed in the EP INRIA 'CLIME' are done via their platform Verdandi.

3.3. Methods for the Applications (in Genetics, Agronomy, Forestry and the Environmental Sciences)

3.3.1. *Theoretical biology: study of specific plant systems or phenomena*

The study of systems as complex as plants requires the development of powerful methodologies of analysis. Models help biologists to explore some specific phenomena. The most obvious way is through simulation. However, mathematical analysis of model behavior or sensitivity analysis are also powerful tools for the diagnosis of biological phenomena. A good example was given by [Mathieu et al., 2008], proving first theoretically the emergence of organogenesis 'rhythms' in plants, before observing them. An ongoing project concerns the model 'NEMA' [Bertheloot et al., 2011], developed jointly with INRA-Grignon (J. Bertheloot, B. Andrieu). The model describes at organ level budgets of both Carbon and Nitrogen in plants. It involves 5 interacting biological functions, with each around 20 parameters. Specific techniques of global sensitivity analysis are developed (PhD of Qiongli Wu) to explore such type of models, with the objective of underlining key biological processes and interactions.

3.3.2. *Risk analysis*

Taking into account the uncertainty in model prediction (uncertainty of parameters, climatic uncertainty), the objective is to quantify for farmers the risk associated to yield. The use of data assimilation (of satellite or aerial images for example) is a crucial point to decrease the level of uncertainty. The important application of such study concerns crop-yield insurance.

3.3.3. *Optimal control of crop cultivation*

How to optimize irrigation or fertilization strategies ? Based on models of plant-soil interactions, we are facing optimal control problems. Our objective is to develop dynamic programming techniques, which seem more adapted to the non-convex situations we are facing. Several questions are of interest: constraints linked to environmental regulations, stochastic control due to climatic uncertainty, control of time-delay systems (due to plant senescence), curse of dimensionality ...

3.3.4. *Optimization of parameters for genetic improvement*

The first step concerns the link of model parameters to genes (or Quantitative Trait Loci) via quantitative genetics model. Then, we can explore through selection process the attainable space of model parameters, in which we can find optima regarding specific criteria (for specific types of climate for example). A new PhD should start soon in collaboration with J. Lecoeur (Syngenta Seeds).

4. Software

4.1. PYGMALION

PyGMAlion (Plant Growth Model Analysis, Identification and Optimization) has become the leading development project in the group. The objective is on one hand to provide modelers with mathematical and statistical tools for model analysis, and on the other hand to capitalize in the same software the different methods developed in the group. The basic idea is that provided the modeler writes his dynamic system of plant growth in a simple frame (defining model state variables, state function, parameters, external inputs, and model observations) then some parameter estimation methods are available, as well as sensitivity analysis, evaluation of criteria for model selection and data assimilation.

5. New Results

5.1. Model Design

5.1.1. *Sensitivity Analysis of Complex Biophysical Models*

5.1.1.1. Background

Sensitivity analysis (SA) is a fundamental tool in the building, use and understanding of mathematical models [44]. Sampling-based approaches to uncertainty and sensitivity analysis are both effective and widely used [38]. For this purpose, Sobol's method is a key one [47]. Since it is based on variance decomposition, the different types of sensitivity indices that it estimates can fulfill different objectives of sensitivity analysis: factor priorization, factor fixing, variance cutting or factor mapping [37]. It is a very informative method but potentially computationally expensive [38]. Besides the first-order effects, Sobol's method also aims at determining the levels of interaction between parameters [48]. In [46], the authors also devised a strategy for sensitivity analysis that could work for correlated input factors, based on the first-order and total-order index from variance decomposition.

5.1.1.2. Algorithm numerical implementation

Computational methods to evaluate Sobol indices sensitivity rely on Monte-Carlo sampling and re-sampling [47], [40]. For k dimensional factor of model uncertainty, the k first-order effects and the 'k' total-order effects are rather expensive to estimate, needing a number of model evaluations strictly depending upon k [43]. Therefore, it is crucial to not only devise efficient computing techniques, in order to make best use of model evaluations [45], but also to have a good control of the estimation accuracy with respect to the number of samples.

With the objective of an efficient computational method for sensitivity analysis of functional-structural tree growth models, we proposed a new estimator based on Homma-Saltelli method to compute Sobol indices, which improves slightly their use of model evaluations thanks to a more balanced resampling strategy. This new estimator can be considered as an effort to improve the efficiency of SA methods for models.

We also derived a theoretical analysis of the error estimation for the sensitivity analysis for the studied class of Sobol's estimators (it can be applied to all the three Sobol's estimators mentioned in this paper) with respect to the sampling size and the number of model evaluations. An analytical test function is used to test the error estimation, and we obtained that the error estimation in this paper gives out a better 'upper bound' than the previous works related to this problem. This error estimation directly relates to the variance of the result, so it can also be used for checking the confidence interval, which is usually difficult to attain.

We derived a theoretical analysis of the error estimation for the sensitivity analysis for the studied class of Sobol's estimators (it can be applied to all the three Sobol's estimators mentioned in this paper) with respect to the sampling size and the number of model evaluations. An analytical test function is used to test the error estimation, and we obtained that the error estimation in this paper gives out a better 'upper bound' than the previous works related to this problem. This error estimation directly relates to the variance of the result, so it can also be used for checking the confidence interval, which is usually difficult to attain.

The result has been accepted to be published in J of Rel. Eng. and Sys. Safety.

Based on the published result for first order index of Sobol's, we also extended this method to the second order index computing which is very important for us to know the precise pairs parameters with interactions between them, with the objective of making best use of the input-output model evaluation matrices that involve in the main part of the sensitivity analysis computing cost. Besides the computing efficiency improvement, one side result we got from the algorithm for second order index is that it can also make the final index has smaller variance so that the result can be more reliable.

Besides Sobol's method, we also tried Morris method to complete the aim of 'screening' parameter of sensitivity analysis.

5.1.1.3. Strategies for FSPM

Global sensitivity analysis (SA) has an important role to play in functional-structural plant growth modeling by assessing the different source of uncertainty help us to gain some insights inner the models so as to explain the behavior of them. Different FSPMs have different scales of model design, which leads to all types of diverse multi-biophysical processes.

To study specifically how global SA can help for FSPMs, SA was applied on a wide variety of functional-structural plant models, typically the 3 FSPMs: firstly a simple source-sink model of maize growth, is used to specifically study the process of carbon (C) allocation among expanding organs during plant growth, with simple plant structure, multi-stage and detailed observations, secondly the GreenLab model of tree growth (applied to poplar tree) characterized by the retroaction of plant functioning on its organogenesis [41], which describes tree structural plasticity in response to trophic competition, lastly a functional-structural model, NEMA [16], describing C and nitrogen (N) acquisition by a wheat plant as well as C and N distributions between plant organs after flowering. This model has the specificity to integrate physiological processes governing N economy within plants: root N uptake is modeled following the transport systems high affinity transport systems (HATS) and low affinity transport systems (LATS), and N is distributed between plant organs according to the turnover of the proteins associated to the photosynthetic apparatus. C assimilation is predicted from the N content of each photosynthetic organ. Consequently, this model is more mechanistic but also more complex than the two previous ones. Another objective is to explore an effective simulation design to help the sensitivity analysis for complicate models with several logically distinct but biological functioning interacted moduls.

All these SA result shadowed a light to the models for us to diagnosis the model behavior and will bring a big step for parameter estimation and experimental simplification in our modelling next.

5.1.2. *A model for Cecropia sciadophylla under fluctuating environmental conditions*

In collaboration with Patrick Heuret (INRA, JRU Ecofog, Kourou, French Guiana), we developed a tree growth model dedicated to *Cecropia sciadophylla*, a neotropical species from the genus Cecropia. These trees have interesting properties from a modeller's point of view: they have a simple architecture, their number of phytomers remain limited even for old individuals, and most importantly, [39] and [49] have developed a methodology based on morphological observations to estimate tree age on C. obtusa and C. sciadophylla respectively. It is therefore possible to fully describe the tree structure and topology from morphological observations, which is very uncommon for trees: for most tree species, their high stature, complex structure, and long life span drastically increase the fieldwork required to collect data at the organ scale and hamper the development, calibration and validation of functional-structural tree growth models and their potentiel applications in the field of forest management.

We used datasets collected on 18 trees in 2007 and 2008 in french Guiana to develop and evaluate our model. Our objective was to analyse the influence of fluctuating environmental conditions on the dynamics of trophic competition within *C. sciadophylla* trees. We defined an integrated environmental factor that includes meteorological medium-frequency variations and a relative index representing the local site conditions for each plant. The meteorological variations were input from pluviometry data, that could be considered as the main fluctuating environmental stress under that tropical climate. The relative index was estimated based on inversion of our model using data from respectively 11 trees for model calibration (those measured in 2007) and 7 trees for model evaluation (those measured in 2008). This study provided a model that can be seen as a tool to disentangle the ontogenic variations (low-frequency trend) and the environmental variations (medium-frequency variation). One paper was accepted for publication [22].

5.1.3. *Using model inversion to analyze the effects of inter-tree competition on four Pine trees grown under two contrasted density conditions*

In collaboration with Guo Hong and Lei Xiangdong (Chinese Academy of Forestry, Beijing, China), we analyzed the characteristics of individual tree response to competition on source-sink balance through the calibration of the GreenLab model.

Four Chinese pine trees (*Pinus tabulaeformis* Carr.) were destructively measured in November 2009 from the nursery garden located in the Yuanyiqi forest farm, Beijing, China. Two 13-year-old trees (T1 and T2) were from a high density plantation (3500N/ha) and two 10-year-old trees (T3 and T4) were from a low density plantation (2000N/ha). We first examined the statistical differences in the tree morphologies and topologies. Significant differences were found for internode diameter, internode biomass and needle biomass between the two densities, but not for internode length. In a second step, we studied the ability of the GreenLab model to simulate the plasticity of pine trees grown under different densities. To fulfil these objectives, it was necessary to find a way to characterize the competition conditions of each tree. Given the inherent difficulty of identifying the most relevant experimental measurements for this characterization, we proposed to represent the effects of competition on the tree growth through a single tree-specific parameter of GreenLab, called characteristic surface area, and to estimate it for each tree by model inversion, together with the more classical endogenous species parameters. This will eventually allow us to examine whether the obtained value of this characteristic surface area could be correlated to other possible indicators of competition pressure. This could pave the way to the development of an individual-based stand growth model including the effects of a competition index.

One paper was submitted to Trees - Structure and Functions.

5.1.4. *Coffee trees and genetics*

In collaboration with Sylvie Sabatier (INRA, AMAP), Philippe de Reffye (CIRAD, AMAP) and Perla Hamon (IRD Montpellier), we studied the architectural and genetic diversities in 5 Coffea species, native from Madagascar. We explore two complementary methods: the genetic diversity using molecular markers (genomic- and/or EST-microsatellites) and the variability of adaptive traits between populations with different ecological niches. We focused on 5 Coffea species endemic to Madagascar, some of which are classified as critically endangered in the World Conservation Union(IUCN) Red List. For each species, architecture and genetic comparative analyses between individuals growing in situ (natural forest) and ex situ (common garden test) are being performed. In parallel, the same populations are analysed using the GreenLab model. These results will be used to study the potential links between the parameters of GreenLab and the allelic distribution in these populations. This is the subject of the PhD of Domohina Andrianasolo (CIRAD, Montpelier and FOFIFA, Antananarivo, Madagascar). This work was presented at the XVIII International Botanical Congress (IBC) [34].

5.1.5. *Methods for tree crown analysis and application to young Eucalyptus*

Based on the pioneer work on coffee trees of Philippe de Reffye, a stochastic model was developed to describe the topological development of trees. In the model, growth and branching processes are driven by the respective probabilities of activity, rest or death of apical and lateral buds. Because of its mathematical formulation, the model inversion can be done analytically Â which is rare Â and parameter values can be estimated from experimental data. The MATLAB softaware GLOUPS developed by Philippe de Reffye was used. We explored the feasibility of calibrating this stochastic model for eucalyptus, which presents the additional difficulty of a continuous growth with no marked endogenous cessation. Incomplete systems were also defined for the case, common with trees, of incomplete datasets. An adequate strategy was defined to sample measurements and applied to five eucalyptus trees (data collected by Pr Lei Xiangdong, Guo Hong and Diao Jun, Chinese Academy of Forestry, Beijing, China).

One paper was accepted for publication [20]

5.2. Model Evaluation and Parameter Estimation

5.2.1. *Maximum Likelihood Estimation*

In [18], a first approach for parameter estimation was introduced based on the assumption of an underlying deterministic model of biomass production and uncorrelated errors in the mass measurements of different types of organs in the plant structure. A novel idea is developed on the modeling plant growh in the framework of non-homogeneous hidden Markov models, for a certain class of plants with known organogenesis (structural development). Unknown parameters of the models are estimated via a stochastic variant of a generalised EM

(Expectation-Maximization) algorithm where both steps (E-step, M-step) are non-explicit. For this reason, the E-step is approximated via a sequential Monte-Carlo procedure (sequential importance sampling with resampling) and the M-step is separated into two steps (Conditional-Maximization), where before applying a numerical maximization procedure (quasi-Newton type), a large subset of unknown parameters is updated explicitly conditioned on the other subset. The model is tested with real data and the results are satisfying. Further work is in progress, including MCMC techniques for parameter estimation (with the collaboration of Dr. Sonia Malefaki from the University of Patras, Greece) and Bayesian type estimation, see [33].

5.2.2. *Convolution Particle Filter for parameter estimation*

Although Kalman filter is applied to various fields and dominated for decades, it is limited by its assumptions of linearity.

Particle filter, which combines Bayesian inference with Monte Carlo sequential sampling approach, is a method using different combinations of random variables sampled directly from the parameter space (or the state space) to estimate parameters and states of a complex system. These combinations, generally called particles, propagate by introducing new observations and provide updated posterior distributions by taking into account their weights. Meanwhile, a resampling procedure is used to prevent the degeneracy problem.

Since classical filtering methods are generally not able to estimate the dynamical state vector along with the unknown parameters, the convolution particle filter is implemented based on convolution kernel approximation to meet the need while modelling with Markovian dynamical system.

Several tests are carried out to examine the performance of the Convolution Particle Filtering method [42], [36], and efforts are made to find the optimal perturbation parameters. The applications of the method rely on the Lotka-Volterra model and the sugar beet model.

Since the quality of the estimations is limited by the number of the observations, the Conditional Iterative Bayesian Filtering method is applied. The principal is simply to use the posteriori distributions as the a priori information to re-perform over and over again the estimation algorithm and each time we introduce the same sequence of observations. This approach helps us to improve significantly the final estimation of the hidden states and the unknown parameters while testing with the virtual data. The bootstrapping method is implemented in order to compare with the results from different methods.

In the case of applications based on dynamical stochastic systems, two types of noise are introduced, one is involved in the modelization technique and the other is attached to the observation procedure. An alternation of deterministic parameter estimation and stochastic parameter estimation is proposed (in progress) which allows us to estimate these two kinds of parameters at the same time.

5.2.3. *Modelling the inter-individual variability of organogenesis in sugar beet populations*

Modelling the inter-individual variability in plant populations is a very important issue to enhance the predictive capacity of plant growth models at the field scale. In the case of sugar beet, this variability is well illustrated by the phyllochron (the thermal time elapsing between the appearance of two successive leaves): if the mean phyllochron remains very stable across seasons, there is a high heterogeneity among individuals. Likewise, seedling emergence may strongly vary within a population, potentially inducing important variations in individual plant productions.

A hierarchical segmented model was used to describe and study the variability of the dynamics of leaf appearance in sugar beet crops. The use of this nonlinear mixed model allows for a better handling of the heterogeneity in the plant population, and gives estimates of this variability: each model's parameter is considered as a random variable, varying from one plant to another around a mean population value, with a given variance.

These mean population values and inter-individual variability can be used as input of functional-structural models, the main issue being then to compute the propagation of these sources of probabilistic uncertainty in the dynamic system of Greenlab.

5.3. Methods for the Applications

5.3.1. *Optimization of Phenotyping based on a Parameter Selection Methodology*

The model Cornflo is a functional plant growth model simulating Corn's growth and yield. Based on it, the classification of environmental scenarios is researched in term of their influences to corn's yield, and their parameter estimation capabilities for Cornflo parameters. The initial qualitative analysis of parameter estimation results shows that environmental scenarios' classification benefit estimation accuracy and identifiability. Currently this project is researched from three aspects. Firstly, different clustering techniques are tested to find the most proper scenarios categories for parameter estimation. Secondly, the scenarios clusters are used for botanical experiments optimization, such as the selection of experimental locations. Lastly, parameter estimation is optimized and researched in a practical use for plant growth models.

5.3.2. *Plant-Soil interactoin and Optimal Control of Irrigation*

This work is performed in collaboration with JC Mailhol (Cemagref). Irrigation scheduling is an important issue for crop management, in a general context of limited water resources and increasing concern about agricultural productivity. Methods to optimize crop irrigation should take into account the impact of water stress on plant growth and the water balance in the plant-soil-atmosphere system. For this purpose, different plant-soil interaction models are proposed to simulate the functional plant growth. In particular, a compartment plant model is designed to integrate water stress impact on different main physiological processes of crop: biomass production, biomass allocation, and foliar senescence. This model is applied and calibrated for maize, in order to predict the harvest index according to the stress undergone by crop during its whole cycle. As for the optimal control problem of irrigation, it can be formulated by considering a price for the crop yield and for the water resource. Dynamic programming is then applied to the plant-soil system to determine an optimal irrigation strategy.

6. Partnerships and Cooperations

6.1. National Initiatives

- Projet FUI, CATIA (2010-2012), Pôle de Compétitivité IAR, Région Centre. Porteur du projet: Nouricia.
- Projet FUI, Aqua Team (2012-2014), Pôle de Compétitivité DREAM, Région Centre. Porteur du projet: Footways.

6.2. International Initiatives

6.2.1. *Visits to International Partners*

- V. Letort, at China Academy of Forestry (April, 2011)
- Y. Chen, at China Academy of Forestry (April, 2011)
- P.-H. Cournède, at Beihang University (April, 2011)
- B. Bayol, at Hanoi Univ. of Science and Technology) (November,December, 2011)

6.2.2. *Visits of International Scientists*

- Sonia Malefaki (Univ. Patras) (July 2011)
- Katarina Smolenova (Univ. Goettingen) (June, September 2011)
- PhuongAnh Nguyen (Hanoi Univ. of Science and Technology) (December, 2011)

6.2.2.1. Internship

- Brenda Delamonica (Smith College, USA) (March - July 2011)
- Octave Etard (Centrale Paris, Imperial College London) (March - July 2011)

7. Bibliography

Major publications by the team in recent years

[1] J. BERTHELOOT, P.-H. COURNÈDE, B. ANDRIEU. *NEMA, a functional-structural model of N economy within wheat culms after flowering: I. Model description*, in "Annals of Botany", 2011, vol. 108, n° 6, p. 1085-1096.

[2] P.-H. COURNÈDE, M. KANG, A. MATHIEU, J.-F. BARCZI, H. YAN, P. DE REFFYE. *Structural factorization of plants to compute their functional and architectural growth*, in "Simulation - Transaction of the Society for modeling and simulation international", July 2006, vol. 82, n° 7, p. 427-438.

[3] P.-H. COURNÈDE, A. MATHIEU, F. HOULLIER, D. BARTHÉLÉMY, P. DE REFFYE. *Computing competition for light in the GreenLab model of plant growth: a contribution to the study of the effects of density on resource acquisition and architectural development*, in "Annals of Botany", 2008, vol. 101, n° 8.

[4] P.-H. COURNÈDE, V. LETORT, A. MATHIEU, M.-Z. KANG, S. LEMAIRE, S. TREVEZAS, F. HOULLIER, P. DE REFFYE. *Some Parameter Estimation Issues in Functional-Structural Plant Modelling*, in "Mathematical Modelling of Natural Phenomena", 2011, vol. 6, n° 2, p. 133-159.

[5] M. KANG, P.-H. COURNÈDE, P. DE REFFYE, D. AUCLAIR, B. HU. *Analytical study of a stochastic plant growth model: application to the GreenLab model*, in "Mathematics and Computers in Simulation", 2008, vol. 78, n° 1, p. 57-75.

[6] V. LE CHEVALIER, M. JAEGER, X. MEI, P.-H. COURNÈDE. *Simulation and visualisation of functional landscapes: effects of the water resource competition between plants*, in "Journal of computer science and technology", Nov. 2007, vol. 22, n° 6, p. 835-845.

[7] V. LETORT, P.-H. COURNÈDE, A. MATHIEU, P. DE REFFYE, T. CONSTANT. *Parametric identification of a functional-structural tree growth model and application to beech trees (Fagus sylvatica)*, in "Functional Plant Biology", 2008, vol. 35(10), p. 951-963.

[8] V. LETORT, P. HEURET, P.-C. ZALAMEA, P. DE REFFYE, E. NICOLINI. *Analysing the effects of local environment on the source-sink balance of Cecropia sciadophylla: a methodological approach based on model inversion.*, in "Annals of Forest Science", 2011.

[9] V. LETORT, P. MAHE, P.-H. COURNÈDE, P. DE REFFYE, B. COURTOIS. *Quantitative Genetics and Functional-Structural Plant Growth Models: Simulation of Quantitative Trait Loci Detection for Model Parameters and Application to Potential Yield Optimization*, in "Annals of Botany", 2008, vol. 101, n° 8.

[10] C. LOI, P.-H. COURNÈDE, J. FRANÇON. *A symbolic method to compute the probability distribution of the number of pattern occurences in random texts generated by stochastic 0L-systems*, in "21st International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'10). DMTCS Proceedings", 2010, vol. AM, p. 473-488.

[11] C. LOI, P.-H. COURNÈDE. *Generating Functions of Stochastic L-Systems and Application to Models of Plant Development*, in "Discrete Mathematics and Theoretical Computer Science Proceedings", 2008, vol. AI, p. 325-338.

[12] A. MATHIEU, P.-H. COURNÈDE, D. BARTHÉLÉMY, P. DE REFFYE. *Rhythms and alternating patterns in plants as emergent properties of a model of interaction between development and functioning*, in "Annals of Botany", 2008, vol. 101, n° 8.

[13] A. MATHIEU, P.-H. COURNÈDE, V. LETORT, D. BARTHÉLÉMY, P. DE REFFYE. *A dynamic model of plant growth with interactions between development and functional mechanisms to study plant structural plasticity related to trophic competition*, in "Annals of Botany", 2009, vol. 103, n° 8, p. 1173-1186.

[14] Q. WU, P.-H. COURNÈDE, A. MATHIEU. *An efficient computational method for global sensitivity analysis and its application to tree growth modelling*, in "Reliability Engineering & System Safety", 2011, vol. In Press:DOI: 10.1016/j.ress.2011.07.001.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[15] C. LOI. *Analyse probabiliste, étude combinatoire et estimation paramétrique pour une classe de modèles de croissance de plantes avec développement stochastique*, Ecole Centrale Paris, May 2011, http://hal.inria.fr/tel-00635011/en.

Articles in International Peer-Reviewed Journal

[16] J. BERTHELOOT, P.-H. COURNÈDE, B. ANDRIEU. *NEMA, a functional-structural model of N economy within wheat culms after flowering: I. Model description*, in "Annals of Botany", 2011, vol. 108, n° 6, p. 1085-1096.

[17] J. BERTHELOOT, Q. WU, P.-H. COURNÈDE, B. ANDRIEU. *NEMA, a functional-structural model of N economy within wheat culms after flowering: II. Evaluation and sensitivity analysis*, in "Annals of Botany", 2011, vol. 107, n° 5, p. 1097-1109.

[18] P.-H. COURNÈDE, V. LETORT, M. Z. KANG, A. MATHIEU, S. LEMAIRE, S. TREVEZAS, F. HOULLIER, P. DE REFFYE. *Some parameter estimation issues in functional-structural plant modelling*, in "Mathematical Modelling of Natural Phenomena", 2011, vol. 6, n° 2, p. 133-159 [*DOI : 10.1051/MMNP/20116205*], http://hal.inria.fr/halsde-00612622/en.

[19] P.-H. COURNÈDE, V. LETORT, A. MATHIEU, M. Z. KANG, S. LEMAIRE, S. TREVEZAS, F. HOULLIER, P. DE REFFYE. *Some Parameter Estimation Issues in Functional-Structural Plant Modelling*, in "Mathematical Modelling of Natural Phenomena", 2011, vol. 6, n° 2, p. 133-159 [*DOI : 10.1051/MMNP/20116205*], http://hal.inria.fr/hal-00631509/en.

[20] J. DIAO, P. DE REFFYE, X. LEI, H. GUO, V. LETORT. *Simulation of the topological development of young eucalyptus using a stochastic model and sampling measurement strategy*, in "Computers and Electronics in Agriculture", 2012, vol. 80, p. 105-114 [*DOI : 10.1016/J.COMPAG.2011.10.019*], http://hal.inria.fr/hal-00651975/en.

[21] A. JULLIEN, A. MATHIEU, J.-M. ALLIRAND, A. PINET, P.-H. COURNÈDE, P. DE REFFYE, B. NEY. *Characterization of the interactions between architecture and source-sink relationships in winter oilseed rape*

(Brassica napus) using the GreenLab model, in "Annals of Botany -London- Oup then Academic Press then Oxford University Press-", 2011, vol. 107, n° 5, p. 765-779 [*DOI :* 10.1093/AOB/MCQ205], http://hal.inria.fr/halsde-00613649/en.

[22] V. LETORT, P. HEURET, P.-C. ZALAMEA, P. DE REFFYE, E. NICOLINI. *Analysing the effects of local environment on the source-sink balance of Cecropia sciadophylla: a methodological approach based on model inversion.*, in "Annals of Forest Science", 2011.

[23] Y. MA, A. M. WUBS, E. HEUVELINK, A. MATHIEU, J. Y. ZHU, B.-G. HU, P.-H. COURNÈDE, P. DE REFFYE. *Simulation of fruit-set and trophic competition and optimization of yield advantages in six Capsicum cultivars using functional-structural plant modelling*, in "Annals of Botany", 2011, vol. 107, n° 5, p. 793-803 [*DOI :* 10.1093/AOB/MCQ223], http://hal.inria.fr/halsde-00613645/en.

[24] B. PALLAS, C. LOI, A. CHRISTOPHE, P.-H. COURNÈDE, J. LECOEUR. *Comparison of three approaches to model grapevine organogenesis in conditions of fluctuating temperature, solar radiation and soil water content*, in "Annals of Botany", 2011, vol. 107, n° 5, p. 729-745 [*DOI :* 10.1093/AOB/MCQ173].

[25] S. TREVEZAS, N. LIMNIOS. *Exact MLE and asymptotic properties for nonparametric semi-Markov models*, in "Journal of Nonparametric Statistics", September 2011, vol. 23, n° 3, p. 719-739 [*DOI :* 10.1080/10485252.2011.555543], http://hal.inria.fr/inria-00625293/en.

[26] F. WANG, M. KANG, Q. LU, V. LETORT, H. HAN, Y. GUO, P. DE REFFYE, B. LI. *A stochastic model of tree architecture and biomass partitioning: application to Mongolian Scots pines*, in "Annals of Botany", 2011, vol. 107, n° 5, p. 781-792.

[27] X. WANG, A. MATHIEU, P.-H. COURNÈDE, A. JULLIEN, J.-M. ALLIRAND, P. DE REFFYE, B. ZHANG. *Variability and regulation of the number of ovules, seeds and pods according to assimilate availability in winter oilseed rape (Brassica napus L.)*, in "Field Crops Research", 2011, vol. 122, n° 1, p. 60-69.

[28] Q.-L. WU, P.-H. COURNÈDE, A. MATHIEU. *An efficient computational method for global sensitivity analysis and its application to tree growth modelling*, in "Reliability Engineering and System Safety", July 2011 [*DOI :* 10.1016/J.RESS.2011.07.001], http://hal.inria.fr/hal-00639539/en.

[29] G. YANG, W. CAO, D. LIU, M. KANG, V. LETORT, M. XING, H. XINYUAN. *Growth and development simulation based on functional-structural model GreenLab for poplar (Salicaceae)*, in "African Journal of Agricultural Research", 2011, vol. 6, n° 13, p. 3071-3077.

International Conferences with Proceedings

[30] C. BAEY, P.-H. COURNÈDE. *Using a hierarchical segmented model to assess the dynamics of leaf appearance in plant populations.*, in "14th Applied Stochastic Models and Data Analysis International Conference (ASMDA 2011)", Rome, Italy, 2011, http://hal.inria.fr/inria-00602419/en.

[31] F. KANG, P.-H. COURNÈDE, J. LECOEUR. *Irrigation Demand Simulation of Large-scale Scenarios and Crop Genotype Selection in Water Shortage Context*, in "Proceedings of the 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011)", 2011.

[32] Z. LI, P.-H. COURNÈDE, J.-C. MAILHOL. *Irrigation Optimization by Modeling of Plant-Soil Interaction*, in "Applied Simulation and Modelling", Crete, Greece, C. ANGELI (editor), ACTA Press, June 22-24 2011.

[33] C. Loi, P.-H. Cournède, S. Trevezas. *Bayesian Estimation in Functional-Structural Plant Models with Stochastic Organogenesis*, in "14th Applied Stochastic Models and Data Analysis International Conference (ASMDA 2011)", Rome, Italy, June 2011, http://hal.inria.fr/hal-00653672/en.

[34] S. Sabatier, D. Andrianasolo, J.-J. Rakotomalala, P. Hamon, P. de Reffye, V. Letort. *Plant architectural and genetic diversities in Coffea native from Madagascar: towards an architecturalfunctional plant growth model applied to Coffea biodiversity perservation*, in "XVIII International botanical congress", 2011, vol. Melbourne Australia, 23-30 july 2011, p. 77-78.

[35] Q.-L. Wu, P.-H. Cournède. *A methodology to study Complex Biophysical Systems with Global Sensitivity Analysis*, in "ESREL2011", Troyes, France, C. Berenguer, A. Grall, C. G. Soares (editors), CRC press, September 2011, http://hal.inria.fr/hal-00639549/en.

References in notes

[36] F. Campillo, V. Rossi. *Convolution Particle Filter for Parameter Estimation in General State-Space Models*, in "IEEE Transactions in Aerospace and Electronics.", 2009, vol. 45, n° 3, p. 1063-1072.

[37] J. Cariboni, D. Gatelli, R. Liska, A. Saltelli. *The role of sensitivity analysis in ecological modelling*, in "Ecological modelling", 2007, vol. 203, p. 167-182.

[38] J. Helton, J. Johnson, C. Salaberry, C. Storlie. *Survey of sampling based methods for uncertainty and sensitivity analysis*, in "Reliability Engineering and System Safety", 2006, vol. 91, p. 1175-1209.

[39] P. Heuret, D. Barthélémy, Y. Guédon, X. Coulmier, J. Tancre. *Synchronization of growth, branching and flowering processes in the south american tropical tree Cecropia obtusa (Cecropiaceae)*, in "American Journal of Botany", 2000, vol. 89, n° 7, p. 1180-1187.

[40] T. Homma, A. Saltelli. *Importance measures in global sensitivity analysis of nonlinear models*, in "Reliability Engineering and System Safety", 1996, vol. 52, p. 1-17.

[41] A. Mathieu, P. Cournède, D. Barthélémy, P. de Reffye. *Rhythms and alternating patterns in plants as emergent properties of a model of interaction between development and functioning*, in "Annals of Botany", 2008, vol. 101, n° 8, p. 1233-1242.

[42] V. Rossi. *Filtrage non linéaire par noyaux de convolution - Application à un procédé de dépollution biologique*, Ecole Nationale Supérieure Agronomique de Montpellier, 2004.

[43] A. Saltelli, Annoni.P, Azzini.I, Campolongo.F, Ratto.M, Tarantola.S. *Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index*, in "Computer Physics Communications", 2010, vol. 181, n° 2, p. 259-270.

[44] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola. *Global Sensitivity Analysis*, The Primer, John Wiley&Sons, 2008.

[45] A. Saltelli. *Making best use of model evaluations to compute sensitivity indices*, in "Computer Physics Communications", 2002, vol. 145, p. 280-297.

[46] A. SALTELLI, S. TARANTOLA. *On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal*, in "Journal of American Statistical Association", 2002, vol. 97, p. 702-709.

[47] I. SOBOL. *Sensitivity analysis for non-linear mathematical models*, in "Mathematical Modeling and Computational Experiment", 1993, vol. 1, p. 407-414.

[48] Q.-L. WU, P.-H. COURNÈDE. *The use of Sensitivity Analysis for the design of Functional Structural Plant Models*, in "Procedia - Social and Behavioral Sciences", 2010, vol. 2, n° 6, p. 7768 - 7769, Sixth International Conference on Sensitivity Analysis of Model Output [*DOI :* DOI: 10.1016/J.SBSPRO.2010.05.219].

[49] P. ZALAMEA, P. STEVENSON, S. MADRIN, P.-M. AUBERT, P. HEURET. *Growth pattern and age determination for Cecropia sciadophylla (Urticaceae)*, in "American Journal of Botany", 2008, vol. 95, p. 263-271.