Proceedings of the 7th International Conference on Functional-Structural Plant Models, Saariselkä, Finland, 9 - 14 June 2013. Eds. Risto Sievänen, Eero Nikinmaa, Christophe Godin, Anna Lintunen & Pekka Nygren. http://www.metla.fi/fspm2013/proceedings. ISBN 978-951-651-408-9.

Reconstruction of leaf area time series using data assimilation on the GreenLab plant growth model and remote sensing

Xing Gong[1,2] , **Thomas Corpetti**[1], **Mengzhen Kang**[2] , **Baogang Hu**[2] , **Laurence Hubert-Moy**[1]
[1]*Costel, Univ Rennes 2, UMR CNRS 6554 LETG, Rennes, France.* [2]*LIAMA, Institute of Automation, Chinese Academy of Sciences, 100080, Beijing, China*
**correspondence:* xgong.nlpr@gmail.com

Highlights: The GreenLab plant growth model is a powerful and complex tool able to simulate plant growth evolution. The calibration of GreenLab model for a specific plant usually requires field experiments with destructive sampling periodically done on individual plants in order to characterize growth and organogenesis. However when one is merely interested on particular aspects of plant growth, such as leaf area evolution or fruit harvest, we demonstrate in our work that a substantially simplified GreenLab model, associated with data assimilation methods, is able to provide satisfactory performances. Based on these techniques, we have successfully applied GreenLab to remote-sensing area where the observation data is highly limited. Experimental results show that qualitative and quantitative criteria on leaf area reconstructed time series are of good quality.

Keywords: GreenLab model, data assimilation, remote-sensing

INTRODUCTION

GreenLab Plant Growth model

The Greenlab plant growth model is an iterative technique that produces complex and variable plant architectures using some model parameters that depend on the kind of plant (Yan et al. 2004). These parameters are generally classified into two categories: those which can be estimated from direct, likely destructive, plant measurements and the hidden parameters deeply implied in the simulation process. The parameters of the first category are connected to the internal structure of the plant and comprise for example its geometry, weight and number of fruits, internodes, leaves, petals, etc. The parameters of the second category include external effects as the photosynthesis efficiency or the biomass demand of various organs.

Practical problems

Obviously, a manual collection of adequate observations and measurements required by Greenlab would be time-consuming and work-intensive. This indeed requires cultivating a plantation of given plant species, to dig out a small amount of individual plant units at a regular time interval and to perform extensive and destructive measurements (Guo Y et al., 2006; Zhan et al., 2003). In a large number of applications, this requirement for observation and measurement is too strong and unlikely to be satisfied.

Core idea: simplify GreenLab, compensating for these simplifications with remote sensing and data assimilation

Fortunately, as is shown in our work, when we are merely interested in certain aspects of plant growth, a simplified Greenlab model with significantly reduced number of parameters can be adequate. With the help of data assimilation methods (Evensen 1994), we are entitled to combine the information provided both by the plant growth model and up-to-date observations.

SHORT PRESENTATION OF GREENLAB

As other plant growth models, the goal of GreenLab is to simulate the growing process by dividing it into a number of consecutive growth cycles (Evers et al. 2007; Yan et al. 2004). In each cycle, depending on influential factors of plant and environmental context, biomass is produced by plant's photosynthesis. After its production, biomass is redistributed among all organs of the plant. The biomass generation formula reads:

$$\mathrm{Q(i)} = \mathrm{E(i)} \cdot S_p \cdot \frac{1}{r}\left(1 - \exp\left(-k \cdot \frac{S(i)}{S_p}\right)\right) \tag{1}$$

In the above equation, $\mathrm{Q(i)}$ stands for biomass generated at i_{th} growth cycle, $\mathrm{E(i)}$ is the environmental factor, S_p is the land area occupation, k is related to plant photosynthesis efficiency, r is a scale parameter, $S(i)$ represents the total leaf area at i_{th} growth cycle. Among above parameters, r, k are species-dependent and have to be determined with learning techniques; $S(i)$ is both an intermediate and final biological index and plays a key role in our application.

Once the biomass is produced, it has to be redistributed among different organs of the plant, when simplifying the model, we adhere to GreenLab's *sink-demand* strategy for biomass partition: all organs of a plant share a common biomass pool generated by photosynthesis (Eq.1). Organs have to compete with each other for available biomass, and their partition lies upon their *demand* ratio of the whole plant. The organ's biomass accumulation $\Delta_{q_0(i)}$ is regarded as its growth.

In our application we are particularly interested in the leaf area evolution of agricultural plants that can be measured from remote sensing images. The iterative transition of leaf area $S(i)$ at i_{th} growth cycle is modeled as:

$$S(i) = \alpha \cdot (S(i-1) - \beta_i) + \Delta_{q_l(i)} \qquad (2)$$

In the above equation, the withering of leaves at i_{th} growth cycle is denoted as β_i. It depends on leaf biological age and life span, and can be easily computed based on GreenLab structural sub-model. $\Delta_{q_l(i)}$ represents leaf's biomass partition at i_{th} growth cycle. The parameter α is in $[0,1]$ and stands for the spoilage in natural growths. It is in practice estimated with learning techniques.

DATA ASSIMILATION

Data assimilation consists in estimating the state of a system by combining different sources of information, namely observation and prior knowledge represented as mathematical dynamical models and statistics. It is a widespread mathematical tool in various research areas such as fluid mechanics, atmosphere physics, oceanology and even pattern recognition (Evensen. 1994).

Among possible solutions to perform data assimilation, we rely in this study on stochastic techniques (Nummiaro et al. 2003, Bonet et al, 1999), mainly because of the non-linearity of the GreenLab model. The overall strategy consists in using two sub-models: i) state transition. ii) correction with observation. On the one hand, many external factors are likely to disturb a theoretical evolution of the plant, and therefore the GreenLab outputs are not guaranteed to fit with observations. On the other hand, using remote sensing, several observations per year can be obtained. This kind of observation is however noisy and even corrupted due to atmosphere disturbance and cloud occultation. In our work, we intend to combine model outputs and observations to reconstruct the mean leaf area evolution over agricultural crops observed from remote sensing.

Because of the complexity of the GreenLab model and the impossibility to assess all the parameters, we decided to simplify the model and to compensate these simplifications with data assimilation strategies. We in addition reformulated the deterministic GreenLab model into a probabilistic one by adding gaussian white noise to biomass production in Eq.1. This probabilistic GreenLab could be regarded as state transition model. The observation model is the satellite imaging process. The combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model, also referred to as PROSAIL, has been used to estimate the Leaf Area Index from remote sensing images (Jacquemoud et al. 2009). In our application, we also mix the PROSAIL with gaussian additive white noise to make it probabilistic and deal with as the observation model.

RESULTS AND DISCUSSION

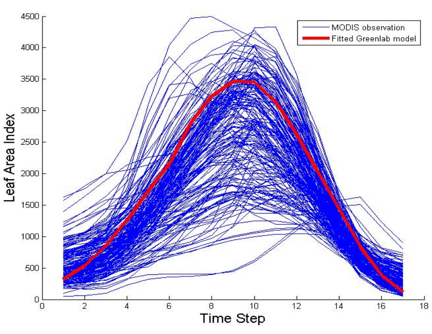

Fig. 1. Leaf area observation sequences of cereal and fitted GreenLab model output. Blue lines are noisy sample sequences calculated with PROSPECT from MODIS data, red line is the output of GreenLab which is fitted with non-linear least square method

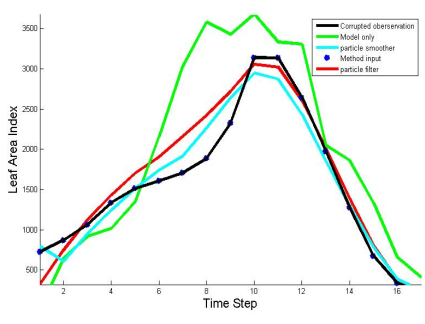

Fig. 2. Result of GreenLab with data assimilation applied to corrupted observation sequence. Dark line are corrupted MODIS observation, red line is the result of GreenLab with particle filter, cyan line from particle smoother, and green line from model only.

The goal of our application is to reconstruct leaf area sequence using GreenLab model with data assimilation methods from medium remote sensing observations, namely temporal series of MODIS images. The area of interested is in Brittany, France. MODIS took a multi-spectral snapshot at a time interval of 10 days, yielding a sequence of 36 observations at each location for a single year. Yet we are only interested on the growth process, which is centered on peak time in summer and extends about 170 days. For this period, leaf area index (LAI) sequences of cereal at various agricultural fields are computed with PROSAIL model and shown with the blue lines in Fig. 1. Because of the atmosphere disturbance and cloud occultation, MODIS data are quite noisy and sometimes even corrupted.

Even though we have significantly reduced the complexity of GreenLab model in our implementation, we still need to estimate three parameters that are r and k from Eq. 1 and α from Eq. 2. We have in practice estimated them automatically using non-linear least square method (Guo et al. 2006, Zhan et al. 2003) on input data. Examples of results are shown in Fig. 1 (a) where the red line is the result of GreenLab with fitted parameters from a set of series.

Tab. 1. Sum of square error (SSE) from different methods at various situation for synthetic data

	Model	**Observe**	**Filter**	**Smoother**	**Sparse**	**Missing**
SSE	346.1	288.6	80.75	70.40.	113.7	293.2

However in practice, because of some problems in the acquisition (aerosols, clouds, etc.), remote sensing data are often noisy and sometimes corrupted. From this noisy & corrupt data, the reconstruction of the LAI is performed using data assimilation. We have tested two variations of stochastic methods: particle filter and particle smoother. Results are shown in Fig. 2 (b). The dark line is a sample of MODIS input data. If we only had the probabilistic GreenLab model (prior knowledge), the green line would be the result we could obtain. After a data assimilation process that combines model (green line) and observations (dark line), the red and cyan lines can be extracted and respectively correspond to the results of particle filter and smoother. From a visual inspection, the data assimilation methods have extracted interesting results. From a quantitative point of view, we have applied our algorithms on synthetic data. The synthetic noisy observation data are generated by adding gaussian noise to a real and uncorrupted observation and by randomly removing 10 points, yielding a time series of 7 points instead of 17. Quantitative results are shown in Tab. 1. As one can observe, the sum of square error (SSE) of corrupted observation is large and this error reduces with assimilation techniques, yielding the particle smoother very efficient.

REFERENCES

Bonet J, Lok T.S. 1999. Variational and momentum preservation aspects of smooth particle hydrodynamic formulations. *Computer Methods in applied mechanics and engineering.* **180(1)**:97-115.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research-All Series,* **99**:10-10.

Evers JB, Vos J, Fournier C, Andrieu B, Chelle M, Struik PC. 2007. An architectural model of spring wheat: evaluation of the effects of population density and shading on model parameterization and performance.. *Ecological Modelling,* **200**:308-320.

Guo Y, Ma YT, Zhan ZG, Li BG, Dingkuhn M, Luquet D, de Reffye P. 2006. Parameter optimization and field validation of the functional-structural model GreenLab for maize. *Annals of Botany* **97**:217-230.

Jacquemoud S, Verhoef W, Baret F, Bacour C, Zarco-Tejada P.J, Asner G.P, François C, Ustin S.L. 2009, PROSPECT + SAIL models: a review of use for vegetation characterization. *Remote Sensing of Environment,* **113**:S56-S66.

Nummiaro K, Koller-Meier E, Van Gool L. 2003. An adaptive color-based particle filter. *Image and Vision Computing,* **91(1)**:99-110.

Yan HP, Kang MZ, de Reffye P, Dingkuhn M. 2004. A dynamic, architectural plant model simulating resource dependent growth. *Annals of Botany* **93**:591-60.

Zhan ZG, Wang YM, de Reffye P, Wang BB, Xiong FL. 2003. Fitting a functional-structural growth model with plant architectural data. In: Hu BG, Jaeger M, eds. *Plant growth modeling and application,* Beijing: Tsing Hua University Press/Springer-Verlag, 108-117.