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.

Global sensitivity analysis of the NEMA model for its parameterization and biological diagnosis

Qiongli Wu [1], **Jessica Bertheloot**[2] and **Paul-Henry Cournède** [1]
[1]*Lab MAS, Ecole Centrale Paris, 92295, Châtenay-Mâlabry Cedex, France*
[2]*INRA,UMR 0462 SAGAH, Beaucouzé Cedex, France*
**correspondence: qiongli.wu@gmail.com*

Highlights: A new comprehensive methodology of global sensitivity analysis is applied to the NEMA model. Besides the common objective of global sensitivity analysis for ranking the importance of individul input factors, we also analyze specifically the functional modules corresponding to the different processes involved in the model, in order to get useful information about the importance and interactions of these physical processes. The results are helpful in the parameterization process and provide new biological insights and diagnosis for NEMA.

Keywords: NEMA, nitrogen, wheat, global sensitivity analysis, SRC, Sobol's method

INTRODUCTION

Functional-structural models of plant growth (FSPM) aim at describing the structural development of individual plants combined with their eco-physiological functioning (photosynthesis, biomass allocation, in interaction with the environment) [Godin and Sinoquet, 2005]. The multi-biophysical processes described in FSPMs and their complex interactions make it difficult to identify the key processes, control variables and parameters driving plant growth. Sensitivity analysis (SA) can help to provide useful insights about the model. First, it helps to simplify parameter estimation by screening the non-influential parameters. Moreover, SA can also provide new biological diagnosis and understanding by the quantitative analysis of interactions between the sub-models corresponding to different processes involved in plant growth.

NEMA (Nitrogen Economy Model within plant Architecture) [Bertheloot et al.,2011] is typically an FSPM with many parameters and many sub-models (modules) describing multi-physical processes in interaction. The main objective of this paper is to perform SA for NEMA to identify the most influential model parameters, processes, and the interactions between them, by taking advantage of several complementary global SA methods.

MATERIALS AND METHODS

From a methodology point of view, there exist different sensitivity analysis methods with different objectives and adapted to different types of models. First of all, the aims of sensitivity analysis need to be considered: Factor prioritization, factor fixing, variance cutting or factor mapping [Saltelli et al.,2008]. Each aim involves different methods. Moreover, the computing cost issue related to the complexity of the models also needs to be considered. For instance, the Standard Regression Coefficients (SRCs) [Cariboni et al., 2007] can be viewed as an interesting trade-off between the accuracy of the analysis and the computing cost, but it is only valid when model's linearity is high. Sobol's method can help us get the interaction information for parameters and functional modules, but the sampling-based Monte-Carlo simulation makes Sobol's method sometimes prohibitive from a computation point of view [Wu et al., 2011]. Especially when we make the module analysis, the interactions between modules are the diverse combinations of interactions between factors within the modules with a lot of higher order Sobol's indices. To circumvent this problem, we propose an adaptation of Sobol's method to group of factors which allows an efficient computation of intra-module and inter-module interactions using the decomposition of total-order indices for group factors associated with sub-modules of the model.

The methodology can be summarized into 5 steps:

1) Non-linearity analysis using SRCs: it gives out the evolution of model non-linearity and a general view of the model dynamic property. This preliminary information is also useful to adapt the following stategy accordingly.

2) Group analysis: it provides the evolution of module importance by computing the evolution of the sensitivity indices of the total set of parameters corresponding to each module of the model. We define the set of factors Ω_m gathering the inputs of one module. Group indices are computed: first order group index

$S^g_{\Omega_m}$ (The main effect of Ω_m) and total order group index $ST^g_{\Omega_m}$ (The total effect of Ω_m). The inter-module interaction is given by $ST^g_{\Omega_m} - S^g_{\Omega_m}$ (it is the sum of interactions of all orders between factor set Ω_m and the other sets).

3) Module by module parameter screening: we compute Sobol's first order index S_i (the main effect of factor X_i, mainly used for 'factor prioritization'), Sobol's total order index ST_i (The total effect of factor X_i, mainly used for 'factor fixing' or screening), $ST_i - S_i$ (The sum of interactions of all orders between factor X_i and the other factors) and Sobol's second order index S_{ij} (The interaction between factors X_i and X_j). We run the SA module by module based on the same sampling points, so that to provide index ST_i of each parameter for screening. For this purpose, we use a time averaging index called TGI [Wu, 2012] to reduce the time dimension of the sensitivity indices.

4) Quantitative intra-module and inter-module interaction analysis: we consider the results from step 2 and step 3 and write the following decomposition $ST^g_{\Omega_m} = \sum_{i\in m} S_i + (S^g_{\Omega_m} - \sum_{i\in m} S_i) + (ST^g_{\Omega_m} - S^g_{\Omega_m})$. $\sum_{i\in m} S_i$ is the main effect brought by all the factors with module *m*, while $S^g_{\Omega_m} - \sum_{i\in m} S_i$ stands for the intra-module interactions and $ST^g_{\Omega_m} - S^g_{\Omega_m}$ is the inter-module interactions as indicated in step 2.

5) Complete analysis for model with the selected parameters.

NEMA [Bertheloot et al.,2009] is a complex functional-structural model describing carbon (C) and nitrogen (N) acquisition by a wheat plant as well as C and N distributions between plant organs after flowering. Nitrogen content of each photosynthetic organ and its remobilization following RubisCO turnover are simulated. The turnover depends on light intercepted and a mobile nitrogen pool, which is enriched by root uptake and nitrogen release from vegetative organs, which is depleted by grain uptake and protein synthesis in vegetative organs. It also accounts for the negative feedback of circulating nitrogen on root uptake, which is formalized following high affinity transport system (HATS) and low affinity transport system (LATS) activities. Organ nitrogen content and light intercepted determine dry matter production by photosynthesis, which is distributed between organs according to their respective demand [Bertheloot et al.,2011]. C assimilation is predicted from the N content of each photosynthetic organ. Inputs of Nitrogen fertilizers are fundamental to get high-yielding crops and a production of high quality with the required protein content. This required a proper understanding of root N uptake regulation and of N determinism on yield and production. Complex interactions exist between root N uptake, N remobilization to grains, and photosynthesis, whose regulatory mechanisms remain far from clear. All the parameters involved for the analysis are given in [Bertheloot et al.,2011].

Basic biological modules are identified for NEMA: Carbon distribution (DMflux), Nitrogen distribution (Nflux), Carbon acquisition via photosynthesis (Photosynthesis), Nitrogen acquisition by roots (RootNuptake), Senescence (TissueDeath). We use subscrips to identify the factors for different plant organs as follows: g for grain, r for root, La for Lamina, Sh for Sheath, In for internode, Pe for Peduncle, Ch for Chaff. Several outputs of interest are considered for both intra-module and inter-module interaction analysis: a) total green area of the plant (AreaGreenTotal), b) total dry mass production of the plant (Production), c) dry mass of the grains (DMgrains), d) Nitrogen mass of the grains (Ngrains) and e) root Nitrogen uptake (RootNuptake). For each output of interest, we did all the 5 steps in the methodology. The evolution of the sensitivity indices are computed for a better investigation of the dynamics of plant growth, but when we aim at screening parameters, we consider the TGI index. The full model involves 83 parameters. We used the parameterization in [Bertheloot et al., 2011].

RESULTS AND DISCUSSION

We select output AreaGreenTotal as a demonstration of how our analysis can provide useful biological insights. The 'non-linearity analysis' and 'group analysis' are shown in Fig.1. Quantitative results for module by module analysis are given directly. A summary of selected parameters for different outputs of interest is also given in Fig.2 to show how 'parameter screening' works for model parameterization.

In Fig.1., with AreaGreenTotal as the output of interest, between 374 ^{o}Cd and 531^{o}Cd, there is one valley for the linearity, during which the lowest linearity appears at 417^{o}Cd to the linearity of 0.48. Such a period in which a strong non-linearity occurs may be characteristic of very specific biological phenomena during plant growth and of high level of interactions between parameters, either known by the modeler or unknown, in which case they should probably be investigated more. The first order sensitivity mainly goes to module DMflux and Nflux, and for a short period to the module Tissuedeath. The interactions between modules

mainly exist for DMflux and Nflux. We also noticed that the module Photosynthesis and RootNuptake contribute little to the variance of the output from the beginning to the end. Moreover, the inter-module interactions for the two modules are also very low at all stages. In module by module analysis, we managed to see these interactions more quantitatively. With TGI in step 4, most intra-module interactions exist in module DMflux, which is as high as 8.75% while the total intra-module interactions for all the five modules is 8.79%. The inter-module interactions between DMflux and Nflux reaches 4.9%.

In Fig.2, 17 factors are selected out of 83 and explain most of the variance of the system. This result has already been used in the parameterization process. The most important factors for module Nfluxe are pretty steady for all the outputs: 1) γ standing for the relative rate of potential grain N filling during cell division, 2) $\sigma_{La}^{N_{ph}}$ standing for relative rate of photosynthetic N synthesis associated to xylem flux for entity Lamina, 3) δ_{La}^{N} standing for relative degradation rates of remobilizable N for entity Lamina. γ rules all the outputs and most of the time has the highest ranking. If we categorize the factors selected in Fig.2 by organs, we can see that in the 4 types of parameters for module DMflux, the ones for organ grain and root tend to have significant effect for all outputs. The parameter of lamina tend to control Nflux, Tissuedeath and Photosynthesis.

The consequences of this study are crucial in several aspects: for parameterization, stressing on which module and on which parameter within each module more care should be taken, but also on whether each module can be parameterized independently (from different experiments for example). Moreover, studying dynamically the interactions between parameters and modules may reveal some biological phenomena of interests, non-visible through simple simulations. In this regards, SA offers new tools in integrative biology.

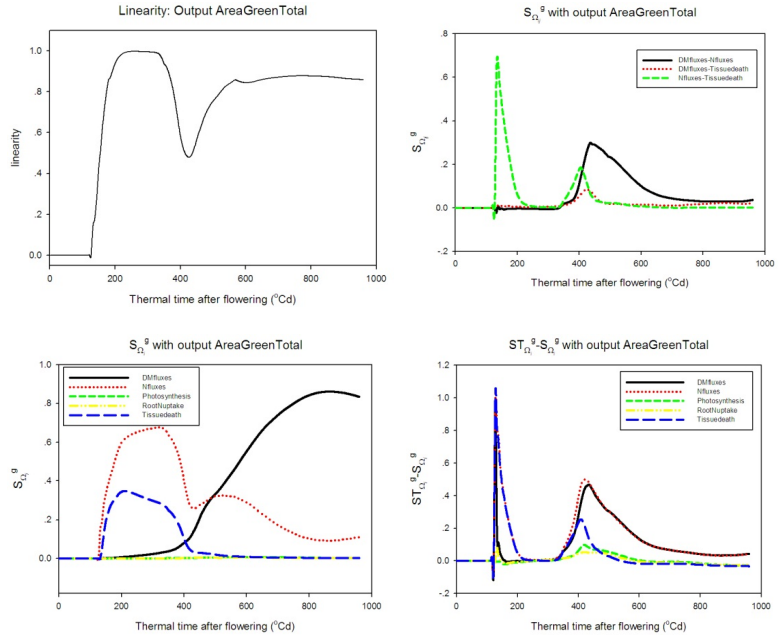

Fig.1. 'non-linearity analysis' and 'group analysis' for output AreaGreenTotal.

S_i ranking	a	b	c	d	e
Module:DMfluxes					
σ_g^M	6	8	4	-	5
α_g	-	12	2	-	8
β_g	21	13	5	-	9
tt_g^{Macc}	9	7	1	7	3
σ_r^M	5	5	-	4	4
α_r	7	6	-	5	6
β_r	3	3	-	3	2
tt_r^{Macc}	1	1	13	2	1
tt_{La}^{Macc}	-	-	8	-	16
tt_{Sh}^{Macc}	-	-	9	-	32
tt_{Ip}^{Macc}	-	-	10	-	17
tt_{Ch}^{Macc}	17	-	-	-	15
Module:Nfluxes					
γ	2	2	11	1	7
$\sigma_{La}^{N_{ph}}$	8	10	7	-	14
δ_{La}^{N}	4	4	6	-	10
Module:Tissuedeath					
d_{La}	11	-	-	-	-
Module:Photosynthesis					
$\omega_{La,2}$	-	9	3	-	11

Fig.2. Selected factor's Si ranking for the complete model for the different outputs

LITERATURE CITED

Godin C, Sinoquet H. 2005. Functional-structural plant modeling. *New Phytologist* **166**:705–708.

Bertheloot J, Cournède P-H, and Andrieu B. 2011. NEMA, a functional–structural model of nitrogen economy within wheat culms after flowering. I. Model description. *Annals of Botany*, **108 (6):** 1085-1096.

Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. 2008. *Global Sensitivity Analysis*. John Wiley&Sons, the primer edition, **24-33**.

Cariboni J, Gatelli D, Liska R, and Saltelli A. 2007. The role of sensitivity analysis in ecological modelling, *Ecological modelling* **203**: 167-182.

Wu Q, Cournède P-H, and Mathieu A. 2012. An efficient computational method for global sensitivity analysis and its application to tree growth modelling. *Reliability Engineering & System Safety* **107**:35-43.

Wu Q. 2012. Sensitivity analysis for functional structural plant modeling, *PhD thesis, Ecole Centrale Des Arts et Manufactures (Ecole Centrale Paris)*, 91-93.