

Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction

David Makowski*, Hervé Monod**, Marie-Hélène Jeuffroy*, Cédric Naud**, Aude Barbottin*

*UMR Agronomie INRA/INA-PG, B.P.01 78850 Thiverval-Grignon, France (Phone : 33 1 30 81 59 92. Fax : 33 1 30 81 54 25. Email : makowski@grignon.inra.fr).

**Unité de biométrie INRA, Domaine de Vilvert, F78352 Jouy-en-Josas Cedex, France (Phone : 33 1 34 65 28 45. Fax : 33 1 34 65 22 17. Email: monod@banian.jouy.inra.fr).

Crop models are complex nonlinear dynamic models simulating several output variables related to crop yield, crop quality, farmer's income and environment. These models are valuable tools for crop management because they can be used to predict the effects of various farmers' practices in function of soil type, climate, and crop characteristics. Crop models can include up to 200 parameters. The estimation of these parameters is an important problem because crop model performances depend for a large part on the accuracy of the parameter estimates. Results obtained with crop models are not reliable when inaccurate parameter values are used. Several methods have been developed recently for estimating crop model parameters from experimental data. However, a large amount of data is always required for estimating accurately crop model parameters. This is particularly true when the model includes genetic parameters. As genetic parameters vary across genotypes, the estimation of these parameters must be based on specific measurements collected for each genotype. Such measurements can be performed in plant breeding experiments but this is very costly and requires a lot of experimental work. Moreover, recent studies have shown that crop model predictions are not always improved when genotypic parameters are estimated genotype per genotype. In this study, we investigate how methods of sensitivity analysis can be used to reduce the quantity of field experiments performed for estimating genetic parameters. The basic principle consists in evaluating the contributions of the genetic parameters to the variance of the model prediction and in estimating genotype per genotype only the key parameters whose uncertainty affects most the outputs.

We consider a crop model developed for winter wheat and named AZODYN (Jeuffroy and Recous, 1999). This model is a dynamic crop model simulating several output variables like, for instance, grain yield, grain protein content, and the amount of residual mineral nitrogen in the soil at harvest. This model is currently used in France for optimizing the dates of nitrogen fertilization and the amounts of applied nitrogen fertilizer. AZODYN includes 70 parameters and 10 of these parameters are supposed to vary across genotypes.

In a first stage, several methods of sensitivity analysis are evaluated and compared with respect to their ability to calculate contribution of genetic and non-genetic parameters to the variance of AZODYN output. Ranges of values are defined for all the parameters on the basis of past experiments, expert knowledge, and literature. These ranges describe the uncertainty about parameter values. Various output variables are considered notably, yield, wheat aerial biomass, nitrogen uptake, grain protein content, and residual soil nitrogen. Several sensitivity indices are defined for each output variable (Homma and Saltelli, 1996 ; Sobol', 2001). We consider first order, second order and total sensitivity indices. The first order sensitivity index associated with a given parameter is defined as the expected fractional reduction in variance

that would be achieved if the parameter were known. The total sensitivity index associated with a given parameter can be thought as the expected fraction of variance that would be left if only this parameter were to stay undetermined. Finally, the second order sensitivity index associated to two parameters is the contribution of the interaction of the two parameters to the total variance. These indices are calculated for the 10 genetic parameters and also for the other parameters of AZODYN. Several Monte Carlo sampling methods are considered successively for calculating sensitivity indices, namely crude Monte Carlo, FAST method, and winding stairs method (Saltelli et al., 2000). Several sample sizes are also considered. The indices obtained for the 10 genetic parameters are compared to the values obtained for the other parameters.

Results of the sensitivity analyses give direct insight into which parameters deserve better genotypic characterization. However, in a second stage of the study, we consider how sensitivity analysis can be more specifically adapted to determining the most relevant genetic parameters. Based on previous expert and experimental knowledge, we define two levels of uncertainty on the genetic parameters, one based on inter-genotypic uncertainty and the other based on intra-genotypic uncertainty. We then evaluate the decrease in output uncertainty when considering intra- rather than inter-variability.

Our study demonstrates that global sensitivity analysis can be used to identify the genetic parameters that must be estimated from plant breeding experiments. The methods considered in this study allow agronomists to determine which subset of parameters accounts for most of the output variance. These methods are useful and easy to interpret. Those factors with a small contribution can be set equal to any value within their range. This contributes to a model simplification and a reduction of the number of experiments performed for estimating crop model parameters. This study provides also information on the performances of the different Monte Carlo sampling methods.

Keywords: crop model, genetic parameters, global sensitivity analysis, Monte Carlo computation, parameter estimation, Sobol' method.

References

- Homma T., Saltelli A. 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52, 1-17.
- Jeuffroy M-H., Recous S. 1999. Azodyn : a simple model simulating the date of nitrogen deficiency for decision support in wheat fertilization. *European Journal of Agronomy*, 10, 129-144.
- Saltelli A., Tarantola S., Campolongo F. 2000. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15, 377-395.
- Sobol' I.M. 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55, 271-280.