GLOBAL SENSITIVITY ANALYSIS FOR CALCULATING THE CONTRIBUTION OF GENETIC PARAMETERS TO THE VARIANCE OF CROP MODEL PREDICTION

David Makowski[a], Cédric Naud[b], Hervé Monod[b], Marie-Hélène Jeuffroy[a], Aude Barbottin[a]

[a] UMR Agronomie INRA/INA-PG, B.P. 01 78850 Thiverval-Grignon, France.
[b] Unité de biométrie INRA, Domaine de Vilvert, F78352 Jouy-en-Josas Cedex, France.
Email: makowski@grignon.inra.fr, monod@banian.jouy.inra.fr

Abstract: Dynamic models are often used to predict the effects of farmers' practices on crop yield, crop quality and environment. These models usually include many parameters that must be estimated from experimental data before practical use. Some of the parameters may vary across genotypes. Such genetic parameters may be estimated from plant breeding experiments but this is very costly and requires a lot of experimental work. Moreover, some of the genetic parameters may account for only a very small part of the output variance and, so, do not deserve an accurate determination. This papers shows how methods of global sensitivity analysis can be used to evaluate the contributions of the genetic parameters to the variance of model prediction. Two methods are applied to a complex nonlinear dynamic crop model for estimating the sensitivity indices associated to 13 genetic parameters. The results show that only 5 genetic parameters have a significant effect on yield and grain quality.

Keywords: Crop model, fast, genetic parameter, global sensitivity analysis, winding stairs

1. INTRODUCTION

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 farmers' practices in function of soil type, climate, and crop characteristics.

Crop models can include up to 200 parameters whose values must be estimated from past experiments. 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. A large amount of data is always required for estimating accurately crop model parameters, in particular 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. This may be due to the small contribution of some of the genetic parameters to the total model output variance.

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. This approach is illustrated below with the AZODYN crop model [1] developed for simulating winter wheat crops. Two methods of global sensitivity analysis are applied to this model in order to evaluate the contribution of 13 genetic parameters to the variances of several output variables of agronomic interests.

2. METHODOLOGY

2.1. The AZODYN model

The AZODYN crop model [1] is a nonlinear dynamic model simulating winter wheat crop in function of numerous input variables describing the characteristics of the crop at the end-of-winter (initial biomass and nitrogen content), soil characteristics (soil texture, organic matter, soil mineral nitrogen), climate (daily radiation and temperature), and nitrogen fertilization (dates and rates of fertilizer applications). In this paper, the input variable are set equal to values obtained in a field located in the north of France (Grignon) and harvested in 2001.

AZODYN is a useful tool for studying the effects of nitrogen management on crop yield, grain quality and risk of pollution by nitrate [2]. The model includes several state variables that are simulated at a daily time step. One of the state variable, namely the nitrogen nutrition index (*NNI*), indicates if the nitrogen content of the crop is at its optimal level or not. AZODYN can be used to predict the characteristics of the crop at harvest, notably grain yield and grain protein content at harvest. Grain yield is an important variable because it determines the farmer's income. Grain protein content is a major grain quality criterion for agro-industries.

Table 1. The 13 genetic parameters of AZODYN and their ranges of variation.

Parameter	Definition	Range	Unit
RDTMAXVAR	Maximal yield	100-137	q.ha^{-1}
Ebmax	Radiation use efficiency	2.7-3.3	g.MJ^{-1}
D	Ratio of leaf area index to critical nitrogen	0.02-0.045	-
REM2	Fraction of remobilized nitrogen	0.5-0.9	-
K	Extinction coefficient	0.6-0.8	-
Eimax	Ratio of intercepted to incident radiation	0.9-0.99	
Tep.flo	Duration between earing and flowering	100-200	°C.day
R	Ratio of total to above ground nitrogen	1.0-1.5	-
P1GMAXVAR	Maximal weight of one grain	47-65	Mg
Lambda	Parameter for calculating nitrogen use efficiency	25-45	-
Mu	Parameter for calculating nitrogen use efficiency	0.6-0.9	-
DJPF	Temperature threshold	150-250	°C.day
NGM2MAXVAR	Maximal grain number	107.95-146.05	-

The AZODYN crop model includes 69 parameters whose values must be estimated before practical use. Among all the parameters, 13 parameters were found to vary across genotypes in past studies. Their values are not perfectly known. The genetic parameters and their ranges of variation are described in table 1. The purpose of this paper is to evaluate the contributions of the 13 genetic parameters to the variances of three model outputs, namely yield, grain protein content and nitrogen nutrition index (*NNI*).

2.2. Comparing different sensitivity analysis methods on AZODYN

2.2.1. Sensitivity indices

We note further Y the output variables of AZODYN. Y will represent in turn yield, grain protein content, and the daily values of *NNI*. Yield and grain protein content are calculated only at harvest, whereas *NNI* is calculated each day between the end-of-winter and harvest. Our approach consists in partitioning the total variance of $V(Y)$ as follows [3]:

$$V(Y) = \sum_{i=1}^{13} V_i + \sum_{i<j} V_{ij} + \sum_{i<j<m} V_{ijm} + V_{1,2...,13}, \qquad (1)$$

where $V(Y)$ is the total variance of the output variable Y induced by the 13 genetic parameters, $V_i = V\left[E(Y|x_i)\right]$ measures the main effect of the parameter x_i, i=1, ..., 13, and the other terms measure the interaction effects. The decomposition (1) is used to derive two types of sensitivity indices defined by

$$S_i = \frac{V_i}{V(Y)}, \qquad (2)$$

$$S_{Ti} = \frac{V(Y) - V_{-i}}{V(Y)}, \qquad (3)$$

where V_{-i} is the sum of all the variance terms that do not include the index *i*. S_i is the *first-order sensitivity index* for the ith parameter. This index measures the main effect of parameter x_i on the output variable Y. S_{Ti} is the *total sensitivity index* for the ith parameter and is the sum of all effects (first- and higher-order) involving the parameter x_i. S_{Ti} takes into account the interactions between the ith parameter and the other 12 parameters. The total sensitivity index can be though as the expected fraction of variance that would be left if only the parameter x_i were to stay undetermined. S_i and S_{Ti} are both in the range (0, 1). The sensitivity indices S_i and S_{Ti} do not differ much from zero when the parameter x_i has a small effect on the output variable Y. On the contrary, if the parameter *i* has a strong effect on Y, the indices take values near from one. The two sensitivity indices S_i and S_{Ti} are equal if the effect of the ith parameter on the model output is independent from the effects of the other parameters.

In the next two sections, we present two methods for estimating the indices (2) and (3) for each parameter and each output variable.

2.2.2. Winding stairs

The calculation of the indices (2) and (3) requires the knowledge of $V(Y)$, V_i, and V_{-i}. The computations can be performed by using a Monte Carlo method [3, 4]. The principle is to generate randomly samples of parameters and to estimate $V(Y)$, V_i, and V_{-i} as follows:

$$\hat{V}(Y) = \frac{1}{N}\sum_{m=1}^{N}\left[f(X_m)\right]^2 - \hat{f}_0^2, \tag{4}$$

where $X_m = (x_{1m},...,x_{im},...,x_{13m})$ is a sample of 13 parameter values drawn in the ranges of variation displayed in table 1, $f(X_m)$ is the simulated value of the output variable, and $\hat{f}_0 = \frac{1}{N}\sum_{m=1}^{N}f(X_m)$.

$$\hat{V}_i = \frac{1}{N}\sum_{m=1}^{N}\left[f\left(X_{(-i)m}^{(1)}, x_{im}\right)f\left(X_{(-i)m}^{(2)}, x_{im}\right)\right] - \hat{f}_0^2, \tag{5}$$

where x_{im} is the mth value of the ith parameter, and $X_{(-i)m}^{(1)}$ and $X_{(-i)m}^{(2)}$ are two different vectors including values of the 12 other parameters defined by $X_{(-i)m}^{(1)} = \left(x_{1m}^{(1)},...,x_{(i-1)m}^{(1)}, x_{(i+1)m}^{(1)}...,x_{13m}^{(1)}\right)$ and $X_{(-i)m}^{(2)} = \left(x_{1m}^{(2)},...,x_{(i-1)m}^{(2)}, x_{(i+1)m}^{(2)}...,x_{13m}^{(2)}\right)$.

$$\hat{V}_{-i} = \frac{1}{N}\sum_{m=1}^{N}\left[f\left(X_{(-i)m}, x_{im}^{(1)}\right)f\left(X_{(-i)m}, x_{im}^{(2)}\right)\right] - \hat{f}_0^2 \tag{6}$$

where $x_{im}^{(1)}$ and $x_{im}^{(2)}$ are two different values of the ith parameter and $X_{(-i)m}$ is a vector including the values of the 12 other parameters.

Different sampling methods can be used to generate the parameter values and organize the computations. Here, we apply the winding stairs sampling scheme [5]. This method was designed to make multiple use of model evaluations. With a single series of N model evaluations, it can compute both the first-order and the total sensitivity indices. The winding stairs method consists in computing the model outputs after each drawing of a new value for an individual parameter. Various procedures can be used to generate the parameter values. Here, the parameter values are generated by Latin hypercube sampling.

The sequence of model outputs generated by the Winding stairs method is shown is table 2. The model outputs are grouped by pairs and are used to compute all the sensitivity indices. For example, the variance (6) is estimated for the first parameter by using the following pairs of model outputs: {1, 2}, {13+1, 13+2}, …, {(N-1)*13+1, (N-1)*13+2}. The variance (5) is estimated for the first parameter by using {2, 13+1}, {13+2, 2*13+1}, …{(N-2)*13+2, (N-1)*13+1}.

The number of model evaluations required for calculating the indices associated to the 13 genetic parameters is equal to 13N (table 2). The winding stairs sampling scheme is applied here with two values of N, specifically N=5000 and N=10000. The Winding stairs method is run 15 times in order to obtain 15 estimates of the first-order and total sensitivity indices for each N value and each parameter. The 15 estimates are averaged and their accuracy is evaluated by computing standard deviation and 95% confidence interval.

Table 2. Sequence of model outputs generated by the Winding stairs method.

Simulation number	Output
1	$f\left(x_{1,1}, x_{2,1}, x_{3,1}, ..., x_{13,1}\right)$
2	$f\left(x_{1,2}, x_{2,1}, x_{3,1}, ..., x_{13,1}\right)$
3	$f\left(x_{1,2}, x_{2,2}, x_{3,1}, ..., x_{13,1}\right)$
4	$f\left(x_{1,2}, x_{2,2}, x_{3,2}, ..., x_{13,1}\right)$
…	
13+1	$f\left(x_{1,2}, x_{2,2}, x_{3,2}, ..., x_{13,2}\right)$
13+2	$f\left(x_{1,3}, x_{2,2}, x_{3,2}, ..., x_{13,2}\right)$
…	
(N-1)*13+1	$f\left(x_{1,N}, x_{2,N}, x_{3,N}, ..., x_{13,N}\right)$
(N-1)*13+2	$f\left(x_{1,1}, x_{2,N}, x_{3,N}, ..., x_{13,N}\right)$
(N-1)*13+3	$f\left(x_{1,1}, x_{2,1}, x_{3,N}, ..., x_{13,N}\right)$
…	
N*13	$f\left(x_{1,1}, x_{2,1}, x_{2,1}, ..., x_{12,1}, x_{13,N}\right)$

2.2.3. Extended FAST

The sensitivity indices are estimated by using a second method named extended FAST [3, 6]. In extended FAST, the sensitivity indices are evaluated by a search curve that scans the space of the 13 parameters, in such a way that each parameter is explored with a selected integer frequency. The basic idea of the method is to convert the 13-dimensional integral in the parameters into a one-dimensional integral by using the transformation function G_i for i=1, …, 13 defined by

$$x_i = G_i\left(\sin \omega_i s\right) \tag{7}$$

where $s \in \,]-\pi, \pi[$ and $\{\omega_i; i = 1,...,13\}$ is a set of integer angular frequencies. The function (7) allows each parameter to be explored globally across its range of variation, as the parameter s is varied over $]-\pi, \pi[$. The implementation of the FAST and extended FAST methods is described in detail in [6].

The method is applied here by using the transformation function $x_i = \dfrac{1}{2} + \dfrac{1}{\pi} \arcsin\left[\sin\left(\omega_i s + \varphi_i\right)\right]$ where φ_i is a random phase-shift parameter drawn in

$[0, 2\pi]$. The frequencies $\{\omega_i; i = 1, ..., 13\}$ are chosen according to the method described by Saltelli [6]. The method requires the computation of several integrals by using a set of model simulations. Here, the indices are calculated from 5000 simulations for each parameter. So, the total number of model simulations is equal to 13*5000. The extended FAST method is run 15 times with different values for the shift parameters in order to derive 15 different estimates of first-order and total sensitivity indices. Like with the Winding stairs method, the 15 estimates are averaged and their accuracy is evaluated by computing standard deviation and 95% confidence interval.

3. RESULT

3.1. Sensitivity indices for yield and grain protein content

Figure 1 shows the total sensitivity indices calculated for the 13 genetic parameters with the Winding stairs method and N=10000. For yield, the parameter *RDTMAXVAR* has the highest total sensitivity index. Its value is equal to 0.77. It means that about 77% of yield variance would be left if only the parameter *RDTMAXVAR* were to stay undetermined. The strong influence of *RDTMAXVAR* is logical because this parameter determined the maximal yield values in the model equations. Two other parameters have a significant influence on yield, namely *Ebmax* and *D*. The total sensitivity indices calculated with the Winding stairs method are equal to 0.18 and 0.17 for these two parameters. The indices of the 10 other parameters are lower than 0.07.

Figure 1 shows that, for grain protein, the parameter with the highest total sensitivity index is *REM2* (index=0.49). This parameter is used by the model to calculate the fraction of the total plant nitrogen that can be allocated to the grains after flowering, and this fraction determines for an important part the value of the grain protein content at harvest. Two other parameters have a total sensitivity index higher than 0.1, namely *R* and *RDTMAXVAR*.
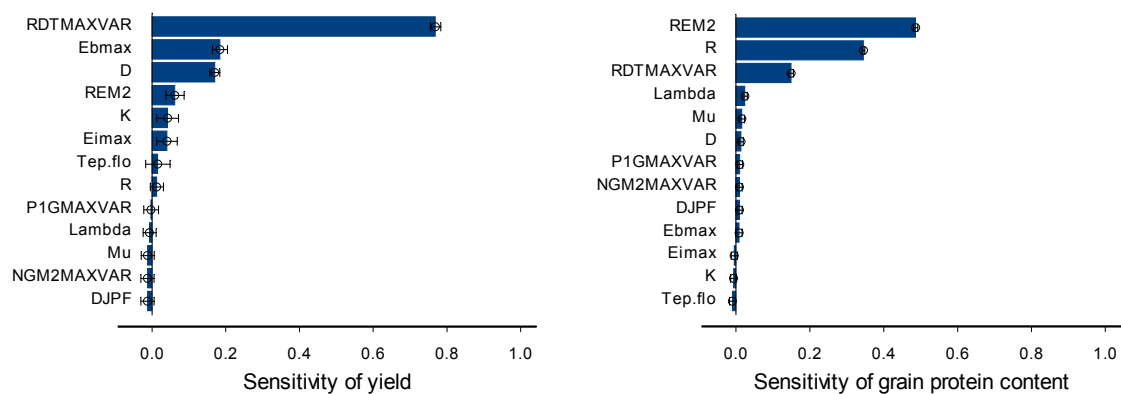

Figure 1. Total sensitivity indices for the output variables yield and grain protein content obtained with the methods Winding stairs (N=10000). The horizontal bars show the average values of the sensitivity indices calculated from 15 estimates for the 13 genetic parameters. Error bars indicate the 95% confidence interval.

Table 3. Sensitivity indices (first order and total effect) for yield and for the 13 genotypic parameters. Mean and standard deviation are calculated from 15 estimates.

| | Winding stairs *N*=10000 | | | | Winding stairs *N*=5000 | | | | Extended Fast | | | |
| | Total sensitivity | | First order sensitivity | | Total sensitivity | | First order sensitivity | | Total sensitivity | | First order sensitivity | |
Parameter	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
RDTMAXVAR	0.769	*0.033*	0.622	*0.036*	0.753	*0.049*	0.627	*0.052*	0.766	*1.2E-02*	0.627	*8.6E-03*
Ebmax	0.184	*0.048*	0.061	*0.045*	0.092	*0.095*	0.155	*0.090*	0.179	*2.6E-03*	0.074	*1.5E-03*
D	0.170	*0.032*	0.055	*0.029*	0.161	*0.053*	0.071	*0.047*	0.159	*8.8E-03*	0.067	*2.8E-03*
REM2	0.062	*0.059*	-0.010	*0.057*	-0.032	*0.122*	0.079	*0.119*	0.043	*1.3E-03*	0.014	*4.7E-04*
K	0.042	*0.069*	-0.010	*0.070*	-0.007	*0.103*	0.050	*0.090*	0.036	*1.5E-03*	0.008	*4.7E-04*
Eimax	0.041	*0.066*	0.010	*0.066*	-0.015	*0.110*	0.075	*0.097*	0.046	*1.5E-03*	0.015	*6.8E-04*
Tep.flo	0.016	*0.078*	-0.015	*0.081*	-0.039	*0.108*	0.050	*0.098*	8.25E-03	*7.3E-04*	1.08E-03	*1.1E-04*
R	0.013	*0.042*	-0.003	*0.039*	0.064	*0.054*	-0.046	*0.052*	1.35E-02	*3.0E-03*	7.14E-04	*1.1E-04*
P1GMAXVAR	-0.003	*0.047*	0.007	*0.041*	0.055	*0.070*	-0.042	*0.070*	1.05E-02	*3.9E-03*	4.17E-04	*1.6E-04*
Lambda	-0.007	*0.042*	0.009	*0.038*	0.057	*0.067*	-0.045	*0.066*	9.58E-03	*2.3E-03*	3.85E-04	*1.6E-04*
Mu	-0.012	*0.042*	0.010	*0.038*	0.052	*0.071*	-0.045	*0.071*	5.50E-03	*1.1E-03*	5.77E-05	*3.5E-05*
DJPF	-0.012	*0.041*	0.008	*0.037*	0.051	*0.067*	-0.046	*0.069*	4.03E-03	*4.6E-04*	4.37E-06	*2.2E-06*
NGM2MAXVAR	-0.012	*0.041*	0.008	*0.037*	0.051	*0.067*	-0.046	*0.069*	4.03E-03	*4.6E-04*	4.37E-06	*2.2E-06*

Table 4. Sensitivity indices (first order and total effect) for grain protein content and for the 13 genotypic parameters. Mean and standard deviation are calculated from 15 estimates.

| | Winding stairs *N*=10000 | | | | Winding stairs *N*=5000 | | | | Extended Fast | | | |
| | Total sensitivity | | First order sensitivity | | Total sensitivity | | First order sensitivity | | Total sensitivity | | First order sensitivity | |
Parameter	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
REM2	0.486	*0.013*	0.480	*0.013*	0.495	*0.024*	0.473	*0.021*	0.493	*5.0E-03*	0.475	*5.1E-03*
R	0.345	*0.009*	0.341	*0.007*	0.339	*0.017*	0.351	*0.013*	0.349	*2.3E-03*	0.337	*2.4E-03*
RDTMAXVAR	0.149	*0.015*	0.111	*0.009*	0.143	*0.016*	0.111	*0.013*	0.144	*2.1E-03*	0.114	*1.6E-03*
Lambda	0.024	*0.014*	0.009	*0.011*	0.017	*0.015*	0.017	*0.012*	0.018	*2.4E-04*	0.012	*4.3E-04*
Mu	0.016	*0.016*	-0.002	*0.011*	0.008	*0.015*	0.007	*0.010*	0.008	*2.5E-04*	0.004	*1.2E-04*
D	0.014	*0.015*	0.013	*0.011*	0.013	*0.013*	0.011	*0.017*	0.024	*7.2E-04*	0.006	*2.8E-04*
P1GMAXVAR	0.011	*0.018*	-0.005	*0.010*	0.004	*0.018*	0.002	*0.011*	2.8E-03	*5.6E-04*	5.4E-05	*2.2E-05*
DJPF	0.010	*0.017*	-0.005	*0.010*	0.003	*0.017*	0.002	*0.011*	2.0E-03	*8.2E-05*	3.7E-07	*1.8E-07*
NGM2MAXVAR	0.010	*0.017*	-0.005	*0.010*	0.003	*0.017*	0.002	*0.011*	2.0E-03	*8.2E-05*	3.7E-07	*1.8E-07*
Ebmax	0.008	*0.020*	0.009	*0.014*	0.021	*0.020*	-0.003	*0.023*	1.9E-02	*4.1E-04*	2.9E-03	*1.2E-04*
Eimax	-0.005	*0.020*	0.006	*0.015*	0.005	*0.023*	-0.002	*0.024*	6.5E-03	*2.1E-04*	5.5E-04	*3.9E-05*
K	-0.007	*0.019*	0.008	*0.015*	0.004	*0.020*	-0.001	*0.019*	6.0E-03	*1.6E-04*	4.1E-04	*3.7E-05*
Tep.flo	-0.010	*0.018*	0.008	*0.016*	0.002	*0.020*	-0.003	*0.018*	2.7E-03	*9.4E-05*	1.7E-04	*1.0E-05*

Tables 3 and 4 show the average values and standard deviations of the first-order and total sensitivity indices obtained for yield and grain protein content with Winding stairs (N=1000 and N=5000) and extended FAST. The average sensitivity indices obtained with the different methods confirm the results shown in figure 1. For yield, the three parameters with the highest sensitivity indices are *RDTMAXVAR*, *Ebmax* and *D*. For grain protein content, the parameters *REM2*, *R* and *RDTMAXVAR* have the highest indices.

The average values of the indices obtained with the different methods are quite similar in most cases (tables 3 and 4). For example, the total sensitivity index obtained for grain protein content and parameter *REM2* is equal to 0.486 with Winding stairs N=10000, to 0.495 with Winding stairs N=5000, and to 0.493 with extended FAST. There are few discrepancies between methods. For example, the parameter with the second highest total sensitivity index is *Ebmax* with Winding stairs N=10000 and extended FAST, but is *D* with Winding stairs N=5000.

The standard deviations of the indices are very different among the methods. With Winding stairs, the standard deviation is lower when the computations are performed from N=10000 simulations than from N=5000. For example, for *REM2* and grain protein content, the standard deviation of the total sensitivity index is equal to 0.013 when N=10000 but is equal to 0.024 when N=5000 (table 4). The use of only 5000 simulations can lead to inaccurate estimations of sensitivity indices. Consequently, with the Winding stairs method, it seems necessary to use at least 10000 simulations to obtain accurate results. Tables 3 and 4 also show that the standard deviations of the indices are much lower with extended FAST than with Winding stairs. For instance, for REM2 and grain protein content, the standard deviation of the estimated values of the total sensitivity index is only equal to 0.005 with extended FAST (table 4). With extended FAST, the variability of the estimated values is due to the variability of the random phase-shift parameter. Note that the results of extended FAST are based only on 13*5000 simulations. Consequently, in terms of efficiency, the extended FAST method seems to perform better than Winding stairs.

Another advantage of the extended FAST method is that it does not give negative estimated values. Negative values are obtained with the Winding stairs method for the parameters that have a very small influence on the model outputs. Of course, such values are unrealistic. For example, for yield and parameter *P1GMAXVAR*, the total sensitivity index is equal to –0.003 with the Winding stairs method and N=10000 (table 3). The value of the same index is equal to 0.001 when computed from the extended FAST method. Negative values could be avoided with the Winding stairs method by calculating the correction term suggested by Homma and Saltelli [4]. But the calculation of this correction term requires additional model evaluations.

The comparison between the first-order and total sensitivity indices allows us to study the contribution of the main effect of the genetic parameters to the total output variances (tables 3 and 4). In most cases, the first-order indices represent an important fraction of the total indices, notably for grain protein content. For example, for *REM2* and grain protein content, the estimated value of the first order index is in the range 0.47-0.48 depending on the method. These values are very near from the average of the 15 estimated values of the total sensitivity index (~0.49).

3.2. Sensitivity indices for *NNI*

The Winding stairs and extended FAST methods were also applied to compute sensitivity indices for the state variable *NNI* (nitrogen nutrition index). This variable is calculated each day between end-of-winter and harvest by the model. Sensitivity indices were computed for each daily value of *NNI*. The results obtained with Winding stairs and extended FAST for parameter *D* (ratio of leaf area to critical nitrogen) are shown in figure 2. At the beginning of the growing period the total sensitivity indices are in the range 0.6-0.8. After day 60, the index decreases sharply and is almost equal to zero after day 100. This result is easily explained by studying the model equations. With AZODYN, *NNI* is calculated in function of the crop biomass and the biomass depends on the leaf area. The leaf area is calculated in function of parameter *D* only at the beginning of the growing period. The leaf area reaches its maximal value after few month of growing. After this stage, the leaf area does not depend on *D* anymore.

As for yield and grain protein content, the results obtained with Winding stairs for *NNI* are inaccurate when *N*=5000; the confidence intervals are larger with *N*=5000 than with *N*=10000. Also, with *N*=5000, the first-order sensitivity indices are higher than the total sensitivity indices at the beginning of the growing period. This is an unrealistic result.
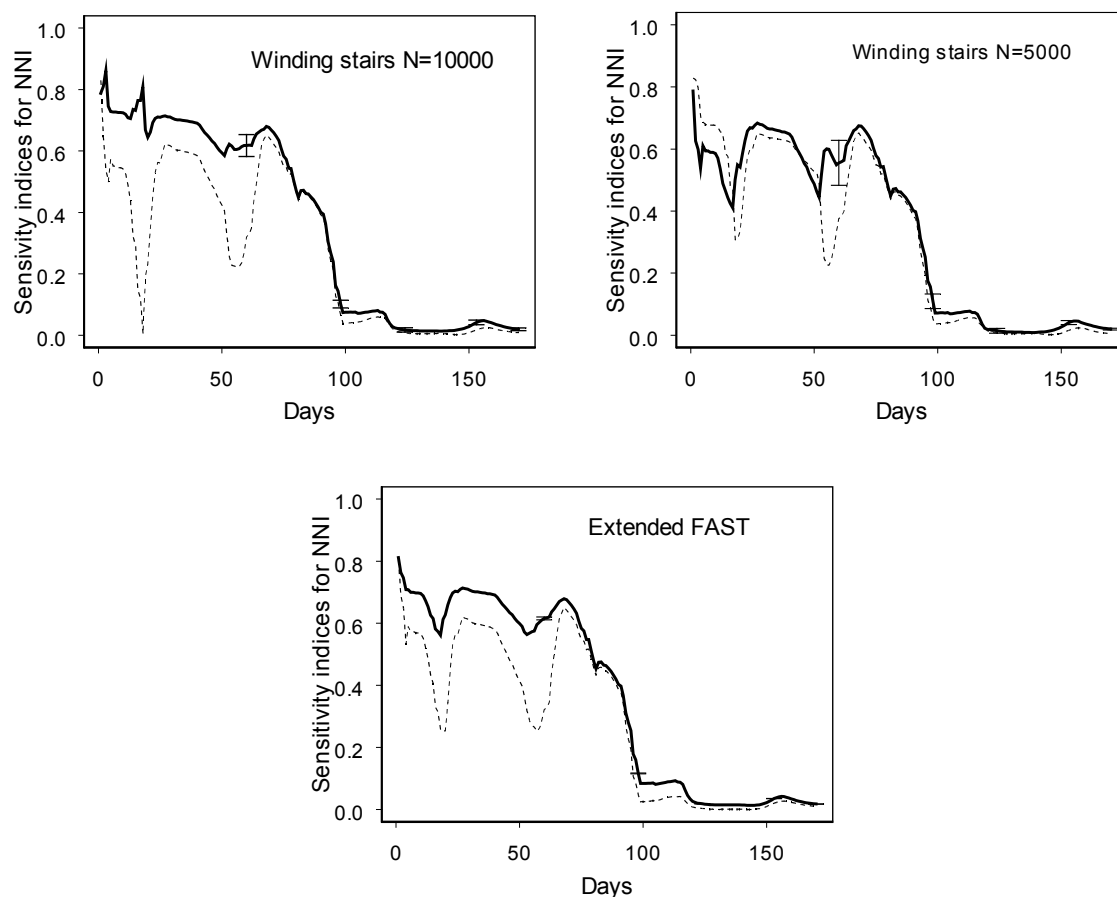

Figure 2. Total sensitivity indices (continuous line) and first-order sensitivity indices (dashed line) for the output variable *NNI* (nitrogen nutrition index) and for parameter *D* (ratio of leaf area to critical nitrogen) obtained with the methods Winding stairs (*N*=10000 and *N*=5000), and Extended FAST. The curves indicate the average values of the sensitivity indices calculated from 15 estimates between the end-of-winter and harvest. The vertical bars indicate the 95% confidence intervals at four dates.

4. CONCLUSION

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.

Our application shows that only 5 parameters have a significant influence on the yield and grain protein content values simulated by the AZODYN crop model. Among these parameters, some can be easily estimated from plant breeding experiments like, for instance, the parameter *RDTMAXVAR* that represents the maximal yield value of a wheat genotype. Others are much more difficult to estimate like the parameter *R* (ratio of total to above ground nitrogen). Note that these results were obtained by running the crop model for a single field and a single year. It would be useful to repeat the analysis for other fields and several years.

The results obtained with the Winding stairs and extended FAST methods are quite similar but the extended FAST method seems to be more efficient. With the Winding stairs method, it is necessary to use at least 10000 model evaluations per parameter for estimating accurately the first-order and total sensitivity indices. With 5000 model evaluations, the Winding stairs method gives inaccurate estimates of sensitivity indices for the parameters that have a small influence on the model outputs.

REFERENCES

1. M-H. Jeuffroy, and S. Recous. Azodyn : a simple model for simulating the date of nitrogen deficiency for decision support in nitrogen fertilization. *European Journal of Agronomy*, 10:129-144, 1999.

2. J-M. Meynard, M. Cerf, L. Guichard, M-H. Jeuffroy, and D. Makowski. Which decision support tools for the environmental management of nitrogen ? *Agronomie*, 22:817-829, 2002.

3. A. Saltelli, S. Tarantola, and F. Campolongo. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 45:377-395, 2000.

4. T. Homma, and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52:1-17, 1996.

5. M.J.W. Jansen, W.A.H. Rossing, and R. Daamen. Monte Carlo estimation of uncertainty contributions from several independent multivariable sources. In J. Gasman and G. van Straten, editors, *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, pages 334-343. Kluwer Academic Publishers, Dordrecht, 1994.

6. K. Chan, S. Tarantola, A. Saltelli, and I. M. Sobol'. Variance-based methods. In Saltelli et al., editors, *Sensitivity Analysis*, pages 167-197. Wiley, New York, 2001.