Environmental Pollution 159 (2011) 3156–3161

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Sensitivity analysis for models of greenhouse gas emissions at farm level. Case study of N_2O emissions simulated by the CERES-EGC model

J.-L. Drouet [a,*], N. Capian [a], J.-L. Fiorelli [b], V. Blanfort [c,d], M. Capitaine [e], S. Duretz [a], B. Gabrielle [a], R. Martin [c], R. Lardy [c], P. Cellier [a], J.-F. Soussana [c]

[a] INRA-AgroParisTech, UMR 1091 Environnement et Grandes Cultures (EGC), F-78850 Thiverval-Grignon, France
[b] INRA, UR 0055 Agro-Systèmes Territoires Ressources (ASTER), F-88500 Mirecourt, France
[c] INRA, UR 0874 Unité de Recherche sur l'Ecosystème Prairial (UREP), F-63100 Clermont-Ferrand, France
[d] CIRAD, Systèmes d'Elevage, F-97387 Kourou, France
[e] ENITA, Agronomie et Fertilité Organique des Sols (AFOS), F-63370 Lempdes, France

A R T I C L E I N F O

Article history:
Received 10 January 2011
Accepted 11 January 2011

Keywords:
Soil–crop model
Farm scale
Greenhouse gas emission
N_2O
Methods of sensitivity analysis

A B S T R A C T

Modelling complex systems such as farms often requires quantification of a large number of input factors. Sensitivity analyses are useful to reduce the number of input factors that are required to be measured or estimated accurately. Three methods of sensitivity analysis (the Morris method, the rank regression and correlation method and the Extended Fourier Amplitude Sensitivity Test method) were compared in the case of the CERES-EGC model applied to crops of a dairy farm. The qualitative Morris method provided a screening of the input factors. The two other quantitative methods were used to investigate more thoroughly the effects of input factors on output variables. Despite differences in terms of concepts and assumptions, the three methods provided similar results. Among the 44 factors under study, N_2O emissions were mainly sensitive to the fraction of N_2O emitted during denitrification, the maximum rate of nitrification, the soil bulk density and the cropland area.

1. Introduction

Agricultural systems, especially livestock systems, are a large source of emissions of greenhouse gases (GHG: N_2O, CO_2, CH_4) and reactive nitrogen (NH_3, NO_x). In 2005, agriculture accounted for 10–12% of total global anthropogenic emissions of greenhouse gases (IPCC, 2007), and ca. 60% of N_2O emissions, and 50% of CH_4 emissions (Soussana et al., 2009). Lifecycle analyses include indirect emissions generated by farm inputs and pre-chain activities. According to this approach, it was estimated that the livestock production systems alone generate directly and indirectly 18% of global GHG emissions as measured in CO_2 equivalents (FAO, 2006). Effective mitigation strategies have to be developed at the farm level (Oenema et al., 2001). Farm models are valuable tools to describe processes, identify interactions between them and test mitigation options (Schils et al., 2007). They are valuable to predict the effects of farmers' practices as a function of soil type, vegetation characteristics and climate. They also aim to predict output variables related to crop yield, pasture quality and farmer's income as well as environmental variables such as GHG emissions and balance. However, farm models are often complex non-linear dynamic models which can include numerous input factors e.g., soil and vegetation parameters, agricultural practices, meteorological input data. To accurately estimate these input factors, a large amount of data is required. The field measurements that provide these data are labour-consuming and costly. Prior to using farm models to assess mitigation options, a sensitivity analysis is required to assess model behaviour and reduce the number of input factors that need to be measured more accurately.

The purpose of this paper is to compare three methods of sensitivity analysis i.e., the Morris method, the rank regression and correlation method and the Extended Fourier Amplitude Sensitivity Test (EFAST) method to analyse their efficiency in reducing the number of input factors to be measured or estimated for parameterization of complex models. This first attempt uses the CERES-EGC model (Gabrielle et al., 2006) which simulates GHG emissions from one component of the farm system i.e., crops. The comparison is drawn up in terms of screening capacity, robustness and computing time. We evaluate the contribution of input factors i.e., soil and vegetation parameters as well as agricultural practices to

* Corresponding author.
E-mail address: Jean-Louis.Drouet@grignon.inra.fr (J.-L. Drouet).

0269-7491/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envpol.2011.01.019

the variance of N_2O emissions. We discuss the results obtained with the three methods and compare our results with those from Lamboni et al. (2009) who carried out a multivariate global sensitivity analysis on the CERES-EGC model.

2. Materials and methods

2.1. Methods of sensitivity analysis

Different options, methods and procedures are available for sensitivity analysis (Saltelli et al., 2008). The most common classifications of available methods distinguish between quantitative and qualitative methods and between local and global techniques (Cariboni et al., 2007). Qualitative methods are aimed at screening, for example, a few active factors within a system with many non-influential ones. They do not give information on the relative difference of importance. Quantitative methods can be designed to give information on the amount of variance explained by each factor. Among qualitative methods, one-at-a-time (OAT) ones estimate the effect of a single factor varying in large ranges when keeping all the others fixed at their nominal values. Global approaches estimate the effect on the output of a factor when all the others are varying, enabling the identification of interactions in non-linear and/or non-additive models. In general, the choice of methods to use is driven by cost, as local or qualitative methods are computationally less expensive.

We used the three methods described by Cariboni et al. (2007) corresponding to three kinds of techniques for global sensitivity analysis. The qualitative Morris method was used to measure the main effect of one factor at a time (OAT), in order to identify input factors which require detailed investigation and identify interactions between input factors. Two quantitative methods were then used to confirm or refute the results using the qualitative Morris method and investigate more thoroughly the effects of input factors on output variables. The first quantitative method was the rank regression and correlation method which is built on regression analysis and can be used to detect the level of linearity of the model (Cariboni et al., 2007). The second quantitative method was based on variance decomposition. Among these variance methods, the Extended Fourier Amplitude Sensitivity Test (EFAST) and the Sobol' methods are the most widely used (Cariboni et al., 2007). We only selected the EFAST method to reduce the number of methods to be compared while exploring different techniques of sensitivity analysis.

2.1.1. The Morris method

The objective of the original elementary effect (EE) method (Morris, 1991) is to determine which input factors may be considered to have effects which are (i) negligible, (ii) linear and additive, or (iii) non-linear or involved in interactions with other factors (Campolongo et al., 2007). The Morris method is a specialized randomized OAT design that has proved to be an efficient and reliable screening technique to identify and rank important variables. It gives a modeller insight into the nature of the influence of input factors on an output of a model with a limited number of model simulations. The method is based on the OAT assumption that if all input factors are changed by the same percentage, the input factor that shows the largest variation in the output is the most sensitive. To perform this, if we consider k input factors $x_1, \ldots, x_k$, the values of each input factor x_i are standardized with their minimal and maximal values to vary in {0,1}. These {0,1} intervals are then discretized in p equispaced values. A sampling strategy generates a multiple number of trajectories through the k-dimensional input factor space i.e., the space over which the factors may vary. Each trajectory provides a single estimation of the EE of the ith input factor, as defined by:

$$\mathrm{EE}_i = [y(x_1, \ldots, x_{i-1}, x_i + \Delta, x_{i+1}, \ldots, x_k) - y(x_1, \ldots, x_k)]/\Delta \quad (1)$$

where Δ is a value in $\{1/(p-1), 2/(p-1), \ldots, 1\}$, p is the number of levels that divide the input factor space, $\{x_1, \ldots, x_k\}$ is the set of input values and y is the model output (see Campolongo et al., 2007; for more details).

Morris (1991) proposed two sensitivity indices, namely the mean (μ) and standard deviation (σ) of the set of EEs for each input factor. μ estimates the overall influence of the factor on the output and σ estimates the ensemble of the higher-order effects of the factor i.e., the non-linear and/or due to interaction effects with other factors.

2.1.2. The rank regression and correlation method

Since regression analysis is based on the linear relationships between the output variable and the input factor, it often performs poorly when this relationship is non-linear, yielding a low value of the R^2 coefficient computed on the raw values. To avoid the problem of non-linearity, rank transformations are frequently employed. The rank transformation method involves replacing the data with their corresponding ranks and therefore assumes that the relationship between the input factor and the output variable is monotonic. A vector of N output values $y = (y_1, \ldots, y_N)$ is generated by repeatedly evaluating the model for a set of N sample vectors $(x_{11}, \ldots x_{1k}), \ldots, (x_{N1}, \ldots, x_{Nk})$, where k is the number of variables. The observations are then replaced by their corresponding ranks 1 (highest value) to N (lowest value). The usual least-squares regression analysis is then performed on the regression equation (see Campolongo et al., 2001; for more details)

Like the Morris method, two sensitivity indices are computed. The standardized rank regression coefficient (SRRC) is based on simple regressions and quantifies the effect of varying each input factor away from its mean. The partial rank correlation coefficient (PRCC) is based on the concepts of correlation and partial correlation and provides a measure of the strength of the monotonic relationship between the input factor and the output variable.

2.1.3. The Extended Fourier Amplitude Sensitivity Test method

The Fourier Amplitude Sensitivity Test (FAST) was developed for uncertainty and sensitivity analysis (Cukier et al., 1973). This method provides a way to estimate the expected value and variance of the output variable as well as the contribution of individual input factors to this variance. Saltelli et al. (1999) developed the Extended FAST (EFAST) method which can also address higher-order interactions between the input factors. The FAST and EFAST methods are independent of any assumptions about the model structure and work for both monotonic and non-monotonic models (Saltelli et al., 2000).

In EFAST, the sensitivity indices are evaluated by a search curve that scans the input factor space in such a way that each factor is explored with a selected integer frequency. The main idea of the method is to convert the k-dimensional integral in the input factors into a one-dimensional integral by using the transformation functions G_i for $i = 1, \ldots, k$, namely

$$x_i = G_i(\sin \omega_i s) \quad (2)$$

where $s \in (-\pi, \pi)$ is a scalar variable and $\{\omega_i\}$ is a set of integer angular frequencies. The G_i function provides a uniformly distributed sample for each factor (see Chan et al., 2001; for more details). The method is applied here by using the transformation proposed by Saltelli et al. (1999):

$$x_i = 1/2 + 1/\pi \arcsin(\sin \omega_i s) \quad (3)$$

Like the Morris and rank methods, two sensitivity indices are computed. The first-order variance provides a measure of the main effect contribution of the input factor to the variance of the output variable. The total variance provides a measure of the total contribution to the output variation due to the input factor i.e., its first-order effect and all higher-order effects due to interactions with other factors.

2.1.4. Implementation of the three methods

The three methods were implemented under the R software package (Venables and Ripley, 2003) version 2.8.0. The Morris method was already implemented under R in a function of the "sensitivity" package. This function was adapted in order to process several output variables at the same time and make it possible to vary simultaneously values of groups of input factors. Five values were used for each input factor: the nominal value of the 'reference farm' and values of ±20% and ±40% from their reference nominal value which define the input factor space. For each input factor, the sensitivity indices μ and σ were computed from 120 trajectories through the input factor space. μ (resp. σ) values were normalized by the sum of the μ (resp. σ) values of all input factors. The rank method and the EFAST method were already implemented under R in the "sensitivity" package and were also adapted to analyse several output variables at the same time. Moreover, the "lhs" package was used when implementing the rank method to re-sample the values of the input factors from Latin hypercube sampling (Iman, 1992). Sensitivity analyses using the rank and EFAST methods were carried out on the 14 input factors which had the largest effect on each output variable using the Morris method (normalized $\mu > 0.001$). Boundary values were ±40% from the nominal value. For each of the 14 input factors, the rank (resp. EFAST) method required 600 (resp. 1000) model simulations.

2.2. The CERES-EGC model

The CERES-EGC model (Fig. 1) was adapted from the CERES suite of soil–crop models (Jones and Kiniry, 1986), with a focus on the simulation of environmental outputs such as nitrate leaching or the emission of nitrogen oxides (Gabrielle et al., 2006). CERES-EGC runs at a daily time step, and requires meteorological and management data as forcing variables and soil and vegetation data as input factors. The nitrous oxide emission module simulates the production of N_2O in soils through both the nitrification and the denitrification pathways (Hénault et al., 2005). Nitrous oxide emissions resulting from both processes are soil-specific proportions of total denitrification and nitrification pathways (see Lehuger et al., 2009; for more details). Soil input factors include soil hydrodynamic properties (e.g., hydraulic conductivity and field capacity), physical properties (e.g., albedo and thermal conductivity), soil texture characteristics (e.g., bulk density and clay fraction) and factors for nitrification and denitrification processes (e.g., maximum rate of nitrification, potential rate of denitrification, fractions of nitrified and denitrified nitrogen).

2.3. The reference farm

The 'reference farm' was a theoretical case study reconstructed from data provided by the French Livestock Institute and description of farming practices within an intensive dairy farm located in North-Eastern France (INRA Mirecourt, Lorraine, 48°17′N, 6°08′E, 300 m above sea level; Fiorelli et al., 2008). It comprised