Amélioration de la qualité de prédiction des variables d'intérêt à partir de l'estimation des paramètres du sol

Article 3 : "The estimation of soil properties using observations on crop biophysical variables and the crop model STICS improves the predictions of agro-environmental variables". Soumis à European Journal of Agronomy..

5.1. Objectif

Nous avons vu dans le chapitre précédent que la quantité d'information apportée par le jeu d'observations, telle qu'on peut la mesurer par analyse de sensibilité détermine la qualité d'estimation des paramètres sol. Nous proposons ici d'étudier comment elle détermine également la qualité des prédictions des variables d'intérêt agroenvironnemental. Les paramètres considérés sont ceux étudiés au chapitre précédent. Les variables retenues, parmi celles étudiées au Chapitre 3, sont les variables agroenvironnementales déterminées à la récolte : rendement *Yld*, qualité de la production *Prot* (teneur en protéine pour le blé) et quantité d'azote minéral dans le sol *Nit*. La bonne prédiction de ces variables permet en effet de raisonner de manière efficace les choix de l'agriculteur vis à vis de son travail technique, en conciliant intérêts agronomique et environnemental. Par exemple, Houlès et al. (2004) ont montré que des cartes de préconisation de doses d'engrais azoté pouvaient être élaborées à partir de prédictions spatialisées de ces variables et de l'optimisation d'un critère agroenvironnemental. Nous étendrons dans cette étude la diversité des jeux d'observations en considérant, comme dans le Chapitre 3, des observations réalisées non seulement sur blé d'hiver mais aussi sur betterave sucrière, culture qui permet d'exprimer davantage les propriétés des sols.

5.2. Méthodes

Nous avons vu au Chapitre 1.3.2, à travers des résultats bibliographiques, que l'amélioration des prédictions à partir de l'estimation repose sur le fait que les variables à prédire et les variables observables ont des sensibilités similaires aux paramètres estimés. Nous essaierons donc d'expliquer l'amélioration des prédictions en fonction de leurs sensibilités aux paramètres du sol et de déterminer les jeux d'observations qui, grâce à leur quantité d'information efficace, permettent de réduire les incertitudes sur les prédictions. Nous avons défini cette réduction comme l'amélioration de la qualité des prédictions issues des valeurs estimées des paramètres relativement à celle des prédictions issues de l'information a priori (sa valeur moyenne). Comme pour le chapitre précédent, l'estimation des paramètres du sol par inversion sera effectuée par la méthode Importance Sampling et la même information a priori sur les paramètres sera considérée (déduite de mesures expérimentales sur le site de Chambry). Dans une première partie, ce travail est réalisé avec des observations synthétiques du couvert végétal, afin d'explorer toutes les configurations d'observations éventuelles. Ces différentes configurations seront ici composées de :

- deux cultures annuelles différentes (blé d'hiver et betterave à sucre),
- quatre climats contrastés caractérisés comme sec, moyen sec, moyen humide et humide (les même que ceux du Chapitre 4),
- deux gammes de profondeurs de sol (de 30 à 100 cm pour les sols peu profonds et de 80 à 160 cm pour les sols profonds),
- trois types/tailles de jeux d'observations : des observations composées de *LAI* seulement, de *LAI*+*QN*, et de *LAI*+*QN*+rendement.

Dans une seconde partie, de vraies observations réalisées sur le bassin versant de Bruyères (voir Chapitre 2.4) seront utilisées afin de valider, de manière réaliste, les résultats obtenus avec les observations synthétiques. Sachant que les sites de Chambry et de Bruyères sont proches et qu'on y retrouve des formations pédologiques voisines (voir Chapitre 2.4), l'information a priori considérée dans l'application aux observations de Bruyères est la même que celle utilisée pour les observations synthétiques (déduite de Chambry).

5.3. Résultats

Précision et amélioration de l'estimation des paramètres

Nous avons vu dans le Chapitre 4.3 les résultats de l'estimation des paramètres, en termes de précision et d'amélioration, lorsque des observations synthétiques sur couvert végétal de blé étaient considérées. Nous allons à présent présenter ces deux quantités dans le cas où des observations synthétiques sur couvert végétal de betterave sont considérées (résultats non présentés dans l'article). Que ce soit en termes de précision (voir le Tableau 5-1) ou en termes d'amélioration (voir le Tableau 5-2), nous voyons que les observations synthétiques de betterave permettent de diminuer significativement le critère, par rapport à ceux obtenus dans le Chapitre 4.3. Les observations de betterave sont donc plus efficaces pour estimer les paramètres du sol. Pour preuve, les paramètres *HCC(1)*, *HCC(2)* et *epc(2)* ont un critère qui diminue énormément grâce à ces observations (sensibilité à ces paramètres plus importante). Par exemple, en forte profondeur de sol, le critère *RE* du paramètre *HCC(1)* passe de 0.81 (observations de blé) à 0.22 (observations de betterave). Nous noterons que la condition initiale *Hinit* est le seul paramètre moins bien estimé qu'avec des observations de blé (sensibilité à ce paramètre moins importante).

	Condition	*argi*	*Norg*	*epc(2)*	*HCC(1)*	*HCC(2)*	*Hinit*	*NO3init*
RRMSE (%)	–	24.3	25	30.5	**7.8**	**15.1**	39	33.7
	+	24.7	25.2	**18.8**	**4.5**	**12.6**	35.1	32.1
	sec	24.2	26.1	25	**6.2**	**13.6**	33.3	31.4
	humide (humide +)	24.8 (24.9)	24.1 (24.6)	24.3 (**18.5**)	6.1 (**4.5**)	14.1 (**11.6**)	40.8 (39.2)	34.4 (32.4)

Tableau 5-1. Précision moyenne d'estimation des paramètres du sol (*RRMSE*) avec des observations synthétiques de betterave, sachant la condition agropédoclimatique : faible profondeur de sol (**–**), forte profondeur de sol (+), climat sec (sec), climat humide (humide) ou climat humide en forte profondeur de sol (humide +). En gras les *RRMSE* inférieurs à 20%.

	Condition	*argi*	*Norg*	*epc(2)*	*HCC(1)*	*HCC(2)*	*Hinit*	*NO3init*
RE	–	0.96	0.86	**0.49**	**0.31**	**0.63**	0.89	0.98
	+	0.90	0.86	0.86	**0.22**	**0.65**	0.83	0.96
	sec	0.96	0.97	**0.68**	**0.24**	**0.66**	**0.74**	0.94
	humide (humide +)	0.90 (0.88)	**0.75** (**0.72**)	**0.68** (0.85)	**0.29** (**0.26**)	**0.62** (**0.57**)	0.98 (0.95)	1 (1)

Tableau 5-2. Amélioration moyenne d'estimation des paramètres du sol (critère *RE*) avec des observations synthétiques de betterave, sachant la condition agropédoclimatique : faible profondeur de sol (**–**), forte profondeur de sol (+), climat sec (sec), climat humide (humide) ou climat humide en forte profondeur de sol (humide +). En gras les *RE* inférieurs à 80%.

La dernière ligne de chacun des deux tableaux précédents, concernant des conditions climatiques humides et une forte profondeur de sol, permet de comparer les résultats de l'estimation des paramètres issus d'observations synthétiques avec ceux issus d'observations réelles sur le bassin de Bruyères. Ces résultats, présentés dans le Tableau 5-3, montrent que les paramètres *HCC(1)* et *epc(2)* sont effectivement estimables avec une bonne précision lorsque des observations sur couvert de betterave sont considérées (*RRMSE* respectivement égal à 7.6 et 18.6%). Nous voyons également que l'estimation de *HCC(1)* est fortement améliorée par l'inversion avec des observations réelles de betterave (*RE* égal à 0.54).

	argi	*Norg*	*epc(2)*	*HCC(1)*	*HCC(2)*	*Hinit*	*NO3init*
RRMSE %	**17.9**	22.5	**18.6**	**7.6**	26.9	55	87.6
RE	0.8	**0.78**	**0.62**	**0.54**	0.95	1.37	1.41

Tableau 5-3. Précision (*RRMSE*) et amélioration (critère *RE*) moyenne d'estimation des paramètres du sol avec des observations réelles de betterave, pour la condition agropédoclimatique : climat humide en forte profondeur de sol (humide +). En gras les *RRMSE* inférieurs à 20% et les *RE* inférieurs à 80%.

Dans le Chapitre 4, nous avons vu que la configuration d'observation avait un effet significatif sur l'amélioration de l'estimation des paramètres : un climat sec et une faible profondeur de sol permettent d'obtenir les meilleures améliorations des paramètres liés à l'état hydrique du sol (*epc(2)*, *HCC(1)*, *HCC(2)* et *Hinit*), les autres étant difficilement estimables. Nous voyons à présent, avec ces nouveaux résultats, que les observations sur couvert végétal de betterave permettent d'améliorer encore plus l'estimation de ces paramètres, mis à part pour la condition initiale *Hinit*. Dans la partie suivante, nous allons montrer comment l'estimation des paramètres, sous différentes configurations d'observations, peut améliorer les prédictions.

Amélioration des prédictions

Les résultats de l'article montrent qu'il est possible d'améliorer la prédiction des variables agronomiques, telles que le rendement et la qualité du rendement, mais qu'il est malheureusement plus difficile d'améliorer celle des variables environnementales, telles que la quantité d'azote encore présent dans le sol à la récolte. Dans le cas des observations synthétiques, les prédictions issues des valeurs moyennes de l'information a priori peuvent être améliorées par l'estimation jusqu'à 61.4% pour le rendement et jusqu'à 58.9% pour la qualité, alors que la teneur en azote du sol ne peut être améliorée que jusqu'à 19.6%. Lorsqu'une amélioration est possible, les résultats montrent que cela vient principalement du fait que les variables à prédire sont sensibles aux mêmes paramètres que le sont les variables observables. De plus, l'article montre qu'il existe un certain degré dans les améliorations possibles, dans le sens où les jeux d'observations acquis dans différentes configurations contiennent des quantités d'information variables permettant d'améliorer l'estimation, et par conséquent la prédiction, de manière plus ou moins significative. Par exemple, nous avons vu que les conditions climatiques dans lesquelles les observations ont été recueillies ont un effet significatif sur l'estimation et par conséquent sur la prédiction, dans le sens où les conditions sèches sont plus efficaces que les conditions humides. Dans le cas des observations synthétiques, les conditions sèches améliorent les prédictions – relativement aux conditions humides – d'environ 25% pour le rendement et la qualité et d'environ 5% seulement pour la teneur en azote du sol. Le type de profondeur de sol a lui aussi un effet important dans le sens où les résultats d'estimation et de prédiction sont de meilleure qualité sur un sol peu profond que sur un sol profond. Toujours dans le cas des observations synthétiques, la prédiction du rendement est d'environ 0.4 fois meilleure pour le rendement, 0.5 fois meilleure pour qualité et 0.1 fois meilleur pour la teneur en azote du sol, lorsqu'un sol peu profond est considéré au lieu d'un sol profond. Pour finir, l'utilisation de jeux d'observations recueillis sur deux différentes cultures (blé et betterave) nous a permis de mettre en évidence l'efficacité des observations sur un couvert de betterave dans le sens où la quantité d'information contenue dans les observations de betterave permet de mieux estimer et de mieux prédire que celle contenue dans les observations de blé. Dans le cas des observations synthétiques, l'observation du couvert de betterave améliore les prédictions – relativement aux observations du blé – d'environ 25% pour le rendement, 10% pour la qualité et 5% pour la teneur en azote du sol. Dans l'application aux observations réelles sur le site de Bruyères, nous voyons qu'il est en effet possible d'améliorer significativement la prédiction du rendement et de la qualité par l'estimation des paramètres du sol : jusqu'à 25% pour le blé ; alors que cela est assez difficile pour la teneur en azote du sol. Pour finir, l'effet de la culture observée révèle son réel potentiel : l'observation de la betterave permet d'améliorer d'environ 22% la prédiction du rendement et de la qualité du blé, par rapport à l'observation du blé lui-même. Les résultats de cet article permettent ainsi de faire un diagnostic assez large des différentes possibilités que l'on a, à partir d'éventuels jeux d'observations, pour estimer les paramètres du sol et pour améliorer les prédictions de variables d'intérêt. D'un point de vue pratique, les résultats de l'article peuvent par exemple permettre d'optimiser le recueil du jeu d'observations, en ne recueillant que celles qui mènent à une estimation et à une prédiction de qualité.

5.4. Article 3 : "The estimation of soil properties using observations on crop biophysical variables and the crop model STICS improves the predictions of agro-environmental variables"

Article soumis à European Journal of Agronomy.

The estimation of soil properties using observations on crop biophysical variables and the crop model STICS improves the predictions of agro-environmental variables.

Hubert Varella*[1], Martine Guérif[1], Samuel Buis[1] and Nicolas Beaudoin[2]

[1] INRA, UMR 1114 INRA-UAPV EMMAH, Domaine Saint Paul, Site Agroparc, F84914 Avignon Cedex 9, France
[2] INRA, Unité d'Agronomie LRM, Rue Fernand Christ, 02007 Laon Cedex, France

* Corresponding author: Hubert Varella

Abstract

The behavior of crops can be predicted when all the parameters of the crop model are well known. Among them, the soil parameters are especially difficult to determine at each location in the study area and they affect the quality of the predictions. Using data observed on crop status in the model is one way of estimating the soil parameters. Nevertheless, the results of parameter estimation depend on the observation set and the results of the predictions are thus also affected. The goal of this study is to assess the value of soil parameter estimation and prediction quality for various observation sets. To achieve it, several observation sets acquired in different conditions (winter wheat and sugar beet crops grown in different weather and cropping conditions) were used to estimate the values of the soil parameters which were then reused in the model to predict the variables of interest. Parameters were estimated using the Importance Sampling method (based on the Bayes theory). The quality of parameter estimation is then calculated (a function of RMSE) as well as the quality of predictions (a function of RMSEP). We worked first with synthetic data and then on real data. The results show that parameters related to soil water content are well estimated and the prediction of the variables of interest can thus be greatly improved. Moreover, parameter estimation and variable prediction are better when the soil is shallow and when the observations are made during dry weather and on sugar beet. The results in this paper can be used to assess the effect of the observation set on the quality of parameter estimation and variable prediction.

Keywords: Bayesian estimation, Importance Sampling, soil parameters, prediction of agro environmental variable, dynamic crop model STICS.

1. Introduction

Crop models are useful tools for simulating or predicting the behavior of crops subjected to different cultural practices. Such predictions are made either on a landscape or a field scale, and are widely used in a lot of agro-environmental work such as crop monitoring, yield prediction or decision making (Gabrielle et al., 2002; Houlès et al., 2004). Crop models can include more than 200 parameters whose values must be estimated from past experiments in order to predict crop behavior (Tremblay and Wallach, 2004; Makowski et al., 2006b). For spatial application a knowledge of soil parameters is even more crucial because they are responsible for much of the variability of the crop model output of variables of interest (Launay and Guérif, 2003). These parameters may be estimated from different techniques: either by soil analysis at different points in the study area, from a soil map and the application of soil transfer functions (Reynolds et al., 2000; Murphy et al., 2003), from remote sensing images (Lagacherie et al., 2008) or by using electrical resistivity measurements (Golovko and Pozdnyakov, 2007). The first method is difficult because of practical limitations, as well as time and financial constraints. Detailed soil maps suited to the scale of precision agriculture or even to that of a catchment are not usually available (King et al., 1994), while the use of remote sensing images or electrical resistivity is still hampered by a lack of robust interpretation of the signal (Lagacherie et al., 2008). Moreover, these techniques do not provide the values of all the soil parameters required for a complex crop model. Fortunately, techniques derived from remote sensing images (Weiss and Baret, 1999) or yield monitoring (Blackmore and Moore, 1999) can provide observations on crop state and thus make it possible to estimate soil parameters through the inversion of crop models.

Studies on the inversion of crop models for accessing soil parameters show that observations on the soil-crop system allow soil parameters to be estimated, and the estimates can reduce the uncertainties associated with the prediction of soil crop variables. Such observations can be made in one or several seasons of different crops. For example, yield maps can be used (Ferreyra et al., 2006) for performing the inversion as well as data derived from remote sensing images (Guérif et al., 2006) or soil water content data (Calmon et al., 1999a), but these studies consider only one type of data from a single crop season. In other studies, some authors have estimated soil parameters by using data from more than one crop season. For example, yield maps made over two crop seasons can be used to estimate parameters (Irmak et al., 2001; Timlin et al., 2001), but in this case also only one type of data was used and not other information provided by observing different crops. Some studies took into account different types of data (Braga and Jones, 2004; Varella et al., 2008). Braga used soil water content data and yield maps for estimating the parameters and compared the effect of both types of data on the quality of parameter estimation and on the prediction of yield and soil water content. It as been proved in this study that estimating soil parameters from soil water content measurements led to better parameter estimates and predictions. Varella estimated the soil parameters using two types of data (derived from remote sensing and yield) and showed that the success of the parameter estimation process depends on the weather and the set of observations available. The sensitivity of the observable variables on the soil parameters is shown to be linked to the quality of the estimate. The weather is shown to be an important factor for the outputs to be sensitive to the soil parameters and especially to those related to the soil water content: the drier the weather, the more sensitive are the outputs to these parameters and the better are their estimates. Finally, too few studies compare the effect of the available observation set in different soil and weather conditions on parameter estimation and on prediction: Braga show that observations on soil water content are better for estimating and predicting than yield maps, but they are very costly and require a lot of experimental work; Varella (2008) shows that dry weather gave better estimates but he only considered synthetic data.

What we propose in this paper is to evaluate the ability of various available observation sets (of different sizes and containing different types of data) and to compare the effect of the number of observed crop seasons (one or two), the type of the observed crop (winter wheat or sugar beet) and cropping conditions (weather and soil depth) on the accuracy of the estimated values and thus on the prediction of the output variables of interest. The crop model used is the dynamic model STICS

(Brisson et al., 2002) and the data consist of crop biophysical variables such as leaf area index (*LAI*), absorbed nitrogen and yield, which are currently available on a farm scale from remote sensing and yield monitors used in precision agriculture. The study is first conducted on a synthetic database designed for generating such variable conditions and as many observations as desired: several observation sets of varying size and several types of observed output variables are then considered on two different crops in different weather and cropping conditions. The synthetic study responds to all the objectives. Using the estimated values of the soil parameters, the prediction performance of agro-environmental variables is then analyzed and compared to that obtained by using the prior information on soil parameter values. Next, the parameter estimation and prediction performance are evaluated on a real experimental database. The results are then compared with those obtained with the synthetic database.

2. Materials and methods

2.1 The crop model, output variables and soil parameters considered

2.1.1 The crop model STICS

The STICS model (Brisson et al., 2002) is a one-dimensional dynamic crop model that simulates water, carbon and nitrogen dynamics in the soil crop system with a daily time step, as well as the behavior of many crops (wheat, sugar beet, corn, peas etc.). It distinguishes several compartments of the crop canopy and segments the soil profile into (at most) 5 layers. It considers the effects of water and nitrogen stress on plant growth and grain yield. The crop is essentially characterized by its above-ground biomass carbon and nitrogen, and leaf area index. The main outputs are agronomic variables (yield, grain protein content) as well as environmental variables (water and nitrate leaching). The STICS model is widely used in a lot of agro-environmental contexts (Ruget et al., 2002; Houlès et al., 2004; Beaudoin et al., 2008).

The model includes more than 200 parameters arranged in three main groups: those related to the characteristics of the plant or to the genotype, those related to the soil and those describing the cropping techniques. In our case, we have used the wheat and sugar beet versions of the model for simulating these crops. The parameters related to the plant genotype were determined from literature, from experiments conducted on specific processes included in the model (e.g. critical nitrogen dilution curve) and from calibrations based on a large experimental database (Flenet et al., 2003; Hadria et al., 2007). The genotype parameter values of both crops used in this study were those given with the model (http://www.avignon.inra.fr/agroclim_stics) and some of them for the wheat crop were previously determined by using a large experimental database. The values of the soil and cultural technique parameters depend on the simulated case. The technique parameters are generally easily established as they are based on farmers' decisions. The soil parameters, in this case, will be estimated for each plot from observations of crops over one or several years.

2.1.2 Output variables considered

In this study, we focus on two types of STICS output variables. The first are observations that can be made on wheat and sugar beet canopies by automated methods. They consist of:

- the leaf area index (LAI_t) and the nitrogen absorbed by the plant (QN_t) at various dates t during the crop season - derived from remote sensing image inversion (Weiss and Baret, 1999; Houborg and Boegh, 2008),
- the yield at harvest (*Yld*) as potentially provided by yield monitoring systems.

These output variables, are called "observable variables".

Secondly, a main objective of this study, apart from the estimation of soil parameter values, is the prediction of some output variables of interest, and its improvement as compared to those obtained without precise soil parameter values. They consist of:

- yield at harvest (*Yld*),
- protein in the grain at harvest (*Prot*) (only computed for the wheat crop),
- nitrogen contained in the soil at harvest (*Nit*).

Yield, grain protein content and nitrogen balance in the soil at harvest are of particular interest for decision making, especially for monitoring nitrogen fertilization

(Houlès et al., 2004). Nitrogen absorbed by the plant is also important to analyze the health and growth of the plant during the crop's growing season (Baret et al., 2006).

2.1.3 The selection of soil parameters to estimate

The complete STICS model contains about 60 soil parameters. For our estimation purposes, it was essential to reduce their number.

First we chose the simplest options for simulating the soil system by ignoring the transfer of water and solutes within soil macropores, as well as capillary rise and nitrification, enabling us to omit the parameters associated with these processes. This choice defines -and limits- the domain of validity of the model considered and hence of the results. It is consistent with the soils contained in the databases we use in this study. In situations where these processes need to be accounted for, new parameters might need to be added to deal with these processes unless an increase in the errors is acceptable. Nevertheless, we have verified on our calibration database that the results found by not considering these processes are very close to those found by considering them: it shows the realism of the assumptions.

A second simplification was to consider the soil as two horizontal layers, each of a given thickness. From the observation of the tillage practices in the region around our reference site, the thickness of the first layer was set at 0.30 m. Based on the measurements made on this precision agriculture experimental site in northern France near Laon, Picardie (Chambry 49.35°N, 3.37°E) (Guérif et al., 2001), we added relations, specific to our conditions, between the initial contents of water *Hinit* and mineral nitrogen *NO3init* of the two soil layers:

$$
\begin{aligned}
Hinit &= Hinit_{first\_layer} = Hinit_{second\_layer} \\
NO3init &= NO3init_{first\_layer} = \frac{2}{3} NO3init_{second\_layer}
\end{aligned}
\tag{1}
$$

Finally, we performed a global sensitivity analysis on the 13 resulting soil parameters (Varella et al., 2008). It allowed us to fix those whose effect on the observed variables was negligible: for each parameter we computed the values of its effects on all the observed variables, and dropped the parameters for which all these values were less than 10% of the total effects generated by the 13 parameters. We thus restricted the study to 7 parameters.

As shown in Table 1, the 7 soil parameters considered characterize both water and nitrogen processes. They refer to permanent characteristics and initial conditions. Among the permanent characteristics, clay and organic nitrogen content of the top layer are involved mainly in organic matter decomposition processes and the soil nitrogen cycle. Water content at field capacity of both layers affects the water (and nitrogen) movements and storage in the soil reservoir. Finally, the thickness of the second layer defines the volume of the reservoir. The initial conditions correspond to the water and nitrogen content, *Hinit* and *NO3init*, at the beginning of the simulation, in this case the sowing date.

Parameter	Definition	Range	Unit
argi	Clay content of the 1^{rst} layer	14-37 (16-25)	%
Norg	Organic nitrogen content of the 1^{rst} layer	0.049-0.131 (0.1-0.13)	%
epc(2)	Thickness of the 2nd layer	0-70 or 50-130* (90-120)	cm
HCC(1)	Water content at field capacity (1^{rst} layer)	14-30 (22.5-29.5)	g g^{-1}
HCC(2)	Water content at field capacity (2nd layer)	14-30 (20-25.5)	g g^{-1}
Hinit	Initial water content (both layers)	4-29 (6-25)	% of weight
NO3init	Initial mineral nitrogen content (1^{rst} layer)	4-86 (6-50)	kg N ha^{-1}

* the first range is for a shallow soil and the second is for a deep soil

Table 1. The 7 soil parameters and their ranges of variation considered for synthetic and real experiments. The ranges of variation of the measured values on the Bruyères database are in brackets.

2.2 Data and numerical experiments

The objectives are to compare the effect of: (i) the size and type of observation set (ii) the number of observed crop seasons, (iii) the type of the observed crop: winter wheat or sugar beet and (iv) the given cropping conditions (weather and soil depth) on the performance of soil parameter estimation and prediction of variables of interest.

2.2.1 Synthetic data and experiments

Data and numerical experiments for parameter estimation

A synthetic database was constituted by first drawing virtual soils and then calculating with the crop model virtual observations of wheat and sugar beet crops grown on these soils in different agricultural and weather conditions. The ranges of variation for parameters (see Table 1) are given from measurements performed on the real agricultural soils of Chambry. These soils are classified as loamy soils in the detailed soil map (Gras et al., 1961).

Two sets of 50 virtual plots with various soil properties were drawn within uniform distributions defined by the ranges given in Table 1: one set is characterized by shallow soils and the other by deep soils, in accordance with the two different ranges considered for the parameter *epc(2)*. One of these 50 virtual plots corresponds to a vector of the 7 soil parameter values, noted θ^{true}.

For each virtual plot and each crop, 8 configurations were chosen for the simulation: 4 contrasting weather patterns and 2 different preceding crops (defining different initial conditions). The weather data were obtained from the meteorological station of Roupy (49.48°N, 3.11°E). The four differ ent datasets, named C1, C2, C3 and C4, correspond to the 1975-1976, 1979-1980, 1972-1973 and 1990-1991 seasons respectively. These are characterized with respect to wheat as a dry season for C1, a medium-dry season for C2, a medium-wet season for C3 and a wet season for C4. The two different preceding crops were peas or sugar beet for wheat crops and a catch crop or bare soil for sugar beet. For each crop considered, the results of parameter estimation and prediction of variables will be shown averaged over the two preceding crops. The cropping techniques used are adapted to the crop considered and the type of soil depth from our knowledge of the farmers' practices in the region: for wheat, the sowing date is set to October 15, the amount of nitrogen fertilizer to 200 kg N per hectare for shallow soils and 240 kg for deep soils; for sugar beet the sowing date is set to March 30, the amount of nitrogen fertilizer to 150 kg N per hectare for shallow soils and 200 kg for deep soils.

We therefore created several sets of synthetic observations of the output variables LAI_t and QN_t at 10 dates t and *Yld* at harvest (i.e. 21 observations) for each virtual plot θ^{true}, crop and weather/preceding crop configuration. The 10 dates were distributed throughout the crop season, with a 30-day interval at the beginning of crop growth and a 15-day interval later on, i.e. November 15, December 12, January 15, February 16, March 15, April 5 and 19, May 3 and 17 and June 07 for wheat; May 29, June 12, 19 and 26, July 6 and 16, August 5 and 25, September 14 and October 04 for sugar beet. Computing the synthetic observations $y_{v,t}$ entailed: (i) simulating with STICS the values of the observable variables at the chosen date and (ii) adding a random error term to the simulated values of the observable output variables. For each virtual plot θ^{true} we calculated:

$$y_{v,t} = f_{v,t}\left(\theta^{true}, x_v\right) + \varepsilon_{v,t} \qquad (2)$$

where $f_{v,t}$ is the STICS prediction for the output *v* at date *t*: *Yld*, LAI_t or QN_t, *t=1,...,10*, x_t is the vector of explanatory variables and $\varepsilon_{v,t}$ is the observation error term. These observation error terms are here assumed to be independent, unbiased and normally distributed. Moreover $\varepsilon_{v,t} \sim N\left(0, \left[\sigma_v^0 f_{v,t}\left(\theta^{true}, x_t\right)\right]^2\right)$ and their magnitude is set from values obtained in field conditions in a precision farming project (Machet et al., 2007; Moulin et al., 2007): σ_{Yld}^0 *=9%*, σ_{LAI}^0 *=17%* and σ_{QN}^0 *=30%*. 1600 sets of synthetic observations for each date and variable *LAI*, *QN* and *Yld* (corresponding to 50 plots x 2 soil depths x 2 crops considered x 8 weather/preceding crop configurations) were thus created.

We did the parameter estimation experiments considering that observations were available for one or two crop seasons, with a possible alternation of crops over the crop seasons. Moreover, to evaluate the sensitivity of the results to the number and type of observations used, several sets of synthetic observations were computed. For one or two observed seasons, the first set (S1) is made up of 10 or 9 observations respectively of *LAI*, the second (S2) is made up of 20 or 19 observations respectively of *LAI* and *QN*, and the third (S3) is made up of 21 observations of *LAI*, *QN* and *Yld* (see Table 2, Wheat and Wheat-Wheat). When only one crop season is considered, the 10 dates defined above for *LAI* and *QN* are used, plus *Yld*; when two crop seasons are considered, alternate dates are used for *LAI* and *QN*, plus *Yld* for the two seasons. Thereby it is possible to compare the effect of the size and type of the observation set for one observed season and also the effect of the number of observed seasons for a given observation set.

Different configurations of parameter estimation experiments were thus conducted by using in turn one of the three observation sets from one or two crop seasons (Table 2). The ability to estimate the parameters will be examined as well as the effect of the size and type of the observation set and the effect of the soil depth and observed crop on parameter estimation. The effect of the available weather data on the soil parameter estimation will finally be studied. For each parameter estimation experiment, quality will be assessed by computing the criterion presented in Section 2.3.3.

Observed crop	Weather	Soil depth	Observation sets
Wheat	C1		S1 = $\{LAI_{t=1, 2, 3, 4, 5, 6, 7, 8, 9, 10}\}$
	C2	–	S2 = S1 + $\{QN_{t=1, 2, 3, 4, 5, 6, 7, 8, 9, 10}\}$
	C3	+	S3 = S2 + $\{Yld_{harvest}\}$
	C4		
Wheat-Wheat	C1-C4		S1 = $\{LAI_{t=3, 5, 7, 9}\}$ + $\{LAI_{t=2, 4, 6, 8, 10}\}$
	C2-C3	–	S2 = S1 + $\{QN_{t=1, 3, 5, 7, 9}\}$ + $\{QN_{t=2, 4, 6, 8, 10}\}$
	C1-C2		S3 = S2 + $\{Yld_{harvest}\}$ + $\{Yld_{harvest}\}$
Sugar beet	C1		S1 = $\{LAI_{t=1', 2', 3', 4', 5', 6', 7', 8', 9', 10'}\}$
	C2	–	S2 = S1 + $\{QN_{t=1', 2', 3', 4', 5', 6', 7', 8', 9', 10'}\}$
	C3	+	S3 = S2 + $\{Yld_{harvest}\}$
	C4		
Wheat-Sugar beet	C1-C4		S1 = $\{LAI_{t=3, 5, 7, 9}\}$ + $\{LAI_{t=1', 3', 5', 7', 9'}\}$
	C2-C3	–	S2 = S1 + $\{QN_{t=1, 3, 5, 7, 9}\}$ + $\{QN_{t=1', 3', 5', 7', 9'}\}$
	C1-C2		S3 = S2 + $\{Yld_{harvest}\}$ + $\{Yld_{harvest}\}$
Sugar beet-Sugar beet	C1-C4		S1 = $\{LAI_{t=2', 4', 6', 8'}\}$ + $\{LAI_{t=1', 3', 5', 7', 9'}\}$
	C2-C3	–	S2 = S1 + $\{QN_{t=2', 4', 6', 8', 10'}\}$ + $\{QN_{t=1', 3', 5', 7', 9'}\}$
	C1-C2		S3 = S2 + $\{Yld_{harvest}\}$ + $\{Yld_{harvest}\}$

The weather pattern C1 corresponds to weather data measured in 1975-1976, C2 to 1990-1991, C3 to 1972-1973 and C4 to 1979-1980. – = shallow soil, + =deep soil. Each of the 3 observation sets (S1, S2 and S3) are considered among the following dates for the wheat: 1) 11/15, 2) 12/12, 3) 01/15, 4) 02/16, 5) 03/15, 6) 04/05, 7) 04/19, 8) 05/03, 9) 05/17, 10) 06/07, and for the sugar beet: 1') 05/29, 2') 06/12, 3') 06/19, 4') 06/26, 5') 07/06, 6') 07/16, 7') 08/05, 8') 08/25, 9') 09/14, 10') 10/04.

Table 2. Description of the different parameter estimation experiments conducted on synthetic observations.

Data and numerical experiments for predictions

Variables of interest are predicted for crop seasons that are independent from those used in the previous step (the creation of observations for the parameter estimation process). For each of the 50 virtual plots defined above and each crop (wheat and sugar beet), 120 configurations for prediction were studied: 2 soil depths, 10 years' weather records, 3 sowing dates (15 days before and after the date used in the previous step), and 4 different cropping techniques (2 fertilizer rates, 20% more or less than the rate used in the previous step, x 2 two preceding crops). The weather data were obtained from the meteorological station of Roupy (49.48°N, 3.11°E) and are different from those used in the parameter estimation process.

For each virtual plot θ^{true}, soil depth and prediction configurations, the values of synthetic observations of the predicted variables of interest (*Yld*, *Prot* and *Nit* at harvest) were simulated with STICS but with no error term added. The same values of the permanent soil properties as those used to create the synthetic observations in the estimation step were used for the simulations, while the values of the initial conditions (*Hinit* and *NO3init*) were randomly generated from the distributions defined in Section 2.1.3. For each parameter estimation experiment defined above, each estimated value of the permanent properties was used to predict the output variables of wheat and sugar beet crops using the STICS model. To make the prediction, two assumptions were made about the initial conditions for the predicted season: either they are unknown and thus fixed as the mean of their prior distribution, or they are known and are thus fixed at their measured values. In this way, it should be possible to quantify the effect of the knowledge of the initial conditions on the quality of the output variable prediction.

Different prediction experiments were done to measure the extent to which the prediction error could be reduced by using the estimated values of the permanent properties. The sensitivity of these results to the size and type of the observation set used for estimating the soil parameters, the soil depth and observed crop is therefore analyzed. The success in reducing the prediction error by using the available weather data for estimating the soil parameters is studied for a single observation season on the wheat crop. Finally, the effect of a possible alternation of crops over two crop seasons is studied on the prediction of the behavior of both crops.

2.2.2 Real experimental data and experiments

Dataset

We used a database based on long-term experiments in the Bruyères catchment (49.52°N, 3.67°E). Soil and crop measurem ents were made from 1991 to 1999 on 36 unreplicated permanent sampling sites which are representative of the main crops and soil types. In the real experiment context, a site represents a plot in the synthetic study. The two main crops in the database are winter wheat and sugar beet, the others being oilseed rape, peas and barley. Weather data were measured continuously either at two locations on the site or from the nearest weather station. Full details on this database can be found in Beaudoin *et al*. (2005; 2008).

In these conditions, observations on crops were much scarcer than in our synthetic database. Of the 36 sites in the Bruyères database, 12 were used which provided enough observations of *LAI*, *QN* and *Yld* (at least 3, 3 and 1 respectively) for at least one crop season and extra observations of variables of interest for at least another season. These sites have similar loamy soils to those considered and characterized in the synthetic database. For these sites, the number of observations of *LAI*, *QN* and *Yld* usable for the parameter estimation experiments and the number of observations of *Yld*, *Prot* and *Nit* usable for the prediction of both crops are presented in Table 3. The number of available sites for parameter estimation experiments varies according to the observed crops: 4 for wheat, 10 for sugar beet and 12 for accumulations of crop seasons. The number of observations of the variable *j* which is predicted for both crops varies from 0 to 8. Unlike the synthetic experiments, the observation dates do not cover the entire crop season and no observations are available at the beginning of the crop season. All the weather patterns involved in the 8-year study of the Bruyères database are characterized as wet and the soil types of the 12 sites used are deep loamy soils.

Observed crop	Site	Observation sets	Prediction sets	
			Wheat predicted	Sugar beet predicted
Wheat	2	{3 *LAI* + 3 *QN* + 1 *Yld*}	{3 *Yld* + 3 *Prot* + 6 *Nit*}	{1 *Yld* + 7 *Nit*}
	5	{3 *LAI* + 3 *QN* + 1 *Yld*}	{3 *Yld* + 3 *Prot* + 5 *Nit*}	{1 *Yld* + 2 *Nit*}
	7	{3 *LAI* + 3 *QN* + 1 *Yld*}	{1 *Yld* + 1 *Prot* + 2 *Nit*}	{1 *Yld* + 2 *Nit*}
	12	{3 *LAI* + 3 *QN* + 1 *Yld*}	{1 *Yld* + 1 *Prot* + 3 *Nit*}	{1 *Yld* + 4 *Nit*}
Sugar beet	1	{3 *LAI* + 3 *QN* + 1 *Yld*}	{1 *Yld* + 1 *Prot* + 3 *Nit*}	{1 *Yld*}
	2	{3 *LAI* + 3 *QN* + 1 *Yld*}	{3 *Yld* + 3 *Prot* + 6 *Nit*}	{1 *Yld* + 7 *Nit*}
	3	{3 *LAI* + 3 *QN* + 1 *Yld*}	{1 *Yld* + 1 *Prot* + 3 *Nit*}	{2 *Yld* + 8 *Nit*}
	4	{3 *LAI* + 3 *QN* + 1 *Yld*}	{3 *Yld* + 3 *Prot* + 8 *Nit*}	{4 *Nit*}
	5	{3 *LAI* + 3 *QN* + 1 *Yld*}	{3 *Yld* + 3 *Prot* + 5 *Nit*}	{1 *Yld* + 2 *Nit*}
	6	{3 *LAI* + 3 *QN* + 1 *Yld*}	{2 *Yld* + 2 *Prot* + 3 *Nit*}	{1 *Yld* + 2 *Nit*}
	7	{3 *LAI* + 3 *QN* + 1 *Yld*}	{1 *Yld* + 1 *Prot* + 2 *Nit*}	{1 *Yld* + 2 *Nit*}
	8	{3 *LAI* + 3 *QN* + 1 *Yld*}	{2 *Yld* + 2 *Prot* + 3 *Nit*}	{ }
	9	{3 *LAI* + 3 *QN* + 1 *Yld*}	{3 *Yld* + 3 *Prot* + 5 *Nit*}	{1 *Yld*}
	10	{3 *LAI* + 3 *QN* + 1 *Yld*}	{2 *Yld* + 2 *Prot* + 5 *Nit*}	{2 *Yld* + 4 *Nit*}
Sugar beet-	1	{3 *LAI* + 9 *QN* + 3 *Yld*}	{1 *Yld* + 1 *Prot* + 3 *Nit*}	{1 *Yld*}
Wheat-Wheat	2	{6 *LAI* + 10 *QN* + 3 *Yld*}	{3 *Yld* + 3 *Prot* + 6 *Nit*}	{1 *Yld* + 7 *Nit*}
	3	{3 *LAI* + 9 *QN* + 3 *Yld*}	{1 *Yld* + 1 *Prot* + 3 *Nit*}	{2 *Yld* + 8 *Nit*}
Sugar beet-	4	{4 *LAI* + 6 *QN* + 2 *Yld*}	{3 *Yld* + 3 *Prot* + 8 *Nit*}	{4 *Nit*}
Wheat	5	{6 *LAI* + 7 *QN* + 2 *Yld*}	{3 *Yld* + 3 *Prot* + 5 *Nit*}	{1 *Yld* + 2 *Nit*}
	6	{3 *LAI* + 7 *QN* + 2 *Yld*}	{2 *Yld* + 2 *Prot* + 3 *Nit*}	{1 *Yld* + 2 *Nit*}
	7	{6 *LAI* + 7 *QN* + 2 *Yld*}	{1 *Yld* + 1 *Prot* + 2 *Nit*}	{1 *Yld* + 2 *Nit*}
	8	{3 *LAI* + 7 *QN* + 2 *Yld*}	{2 *Yld* + 2 *Prot* + 3 *Nit*}	{ }
	9	{3 *LAI* + 7 *QN* + 2 *Yld*}	{3 *Yld* + 3 *Prot* + 5 *Nit*}	{1 *Yld* }
	10	{3 *LAI* + 6 *QN* + 2 *Yld*}	{2 *Yld* + 2 *Prot* + 5 *Nit*}	{2 *Yld* + 4 *Nit*}
Wheat-Wheat	11	{7 *QN* + 2 *Yld*}	{2 *Yld* + 2 *Prot* + 8 *Nit*}	{2 *Yld* + 8 *Nit*}
Wheat	12	{3 *LAI* + 3 *QN* + 1 *Yld*}	{1 *Yld* + 1 *Prot* + 3 *Nit*}	{1 *Yld* + 4 *Nit*}

Table 3. Description of the different parameter estimation experiments and the predictions conducted on the selected data of Bruyères database, showing the available observations in the observation and prediction sets.

Numerical experiments

Similarly to the study on synthetic data, the parameter estimation experiments done on this real database depended on the available observations. As shown in Table 3, parameter estimation involves alternating observations on wheat and sugar beet and accumulated observations on wheat and sugar beet over two and three crop seasons. As for the synthetic experiments, the variables of interest were predicted by using the estimated values of the permanent soil parameters and by using the mean value of the distribution of the initial conditions or their measured values. We assume for the experiments on real data that the true values θ^{true} of the parameters and of the observations of the predicted variables, denoted below $f_j\left(\theta^{true}\right)$, are those measured on the sites. As shown in Tab. 1, the true values θ^{true} vary between 16 and 25 for *argi*, 0.1 and 0.13 for *Norg*, 90 and 120 for *epc(2)*, 22.5 and 29.5 for *HCC(1)*, 20 and 25.5 for *HCC(2)*, 6 and 25 for *HCC(1)* and 6 and 50 for *NO3init*.

2.3 Parameter estimation

2.3.1 Method of parameter estimation

We chose a Bayesian method which takes account of existing information on the parameters to be estimated (this improves the quality of the estimation) and computes an estimate of the posterior probability distribution of parameter values (Makowski et al., 2002; Gaucherel et al., 2008).

The posterior parameter distribution is given by Bayes' theorem:

$$\pi(\theta / Y) = \frac{\pi(Y / \theta)\pi(\theta)}{\pi(Z)} \tag{3}$$

where Y is the vector of total observations of size K, $\pi(\theta / Y)$ is the posterior parameter distribution, $\pi(\theta)$ is the prior parameter distribution, $\pi(Y)$ is a constant of proportionality determined by the requirement that the integral of $\pi(\theta / Y)$ over the parameter space equals 1, and $\pi(Y / \theta)$ is the likelihood function. The likelihood is the probability of the data Y given the parameters θ. Its value is determined from the probability distribution of the errors of modeled and observed data. It is readily seen that both the prior distribution and the new data affect the posterior parameter distribution.

The two most popular families of bayesian methods are Importance Sampling and MCMC (Gilks et al., 1995). These methods are based on Monte Carlo simulations and thus require a large number of model evaluations. The Importance

Sampling method (Beven and Freer, 2001; Makowski et al., 2002) was chosen for our study.

The principle of the Importance Sampling method (Beven and Binley, 1992; Beven and Freer, 2001) is to approximate the posterior parameter distribution $\pi(\theta / Y)$ given in (3) by a discrete probability distribution (θ_n, p_n), *n=1,...,N*, $\sum_{n=1}^{N} p_n = 1$, where p_n is the probability associated with the parameter vector θ_n. In our case, the method proceeds as follows:

(1) Randomly generate *N* vectors θ_n, *n=1,...,N*, from the prior parameter distribution $\pi(\theta)$.

(2) Calculate the likelihood values $\pi(Y / \theta_n)$ for *n=1,...,N*, associated with the different generated parameter vectors.

(3) Calculate $p_n = \dfrac{\pi(Y / \theta_n)}{\sum_{m=1}^{N} \pi(Y / \theta_m)}$

The pairs (θ_n, p_n), *n=1,...,N*, can be used to determine various characteristics of the posterior distribution, including the mean of the posterior joint distribution of θ, $\bar{\theta}^{post} = \sum_{n=1}^{N} p_n \theta_n$.

In this study, we assume that the errors of simulated and observed data are independent between dates and variables and follow normal distributions of zero mean and standard deviation σ_k. Thus, we use the following likelihood function:

$$\pi(Y / \theta) = \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{ -\frac{1}{2\sigma_k^2} \left[y_k - f_k(\theta, x) \right]^2 \right\} \qquad (4)$$

The parameters are assumed to be independent in our case. The prior distribution $\pi(\theta)$ is thus the product of the different marginal prior distributions. We have implemented the Importance Sampling method using Matlab® software, and the application of the procedure simply requires the definition of the total number of generated parameter vectors *N*. A preliminary study of the convergence of the estimates allowed us to set this number at 10 000.

2.3.2 Prior information

As prior information on the parameter values, we used ranges of variation defined in Table 1 and assumed uniform (i.e.: non-informative) distributions. These ranges were obtained from measurements made on precision agriculture site of Chambry (see Section 2.1.3) on loamy soils. The soils represented in the Bruyères database are deep loamy soils and are thus similar to those of the experimental fields of Chambry, but deeper. Moreover, the prior distribution determined on Chambry is convenient to describe the parameter variability observed at Bruyères, as shown in Tab. 1. This prior distribution is thus used to estimate parameters on both experiments (synthetic and real).

2.3.3 Criterion expressing the quality of parameter estimation

We proposed a criterion, noted RE_i (for Relative Error of the parameter *i*), to quantify the quality of parameter estimation. It computes, for each parameter, the ratio of the error of the estimated parameter $\overline{\theta}_i^{post}$ to that of the prior information $\overline{\theta}_i^{prior}$:

$$RE_i = \frac{RMSE(\overline{\theta}_i^{post})}{RMSE(\overline{\theta}_i^{prior})} \tag{5}$$

$RMSE\left(\overline{\theta}_i^{post}\right) = \sqrt{\frac{1}{P}\sum_{p=1}^{P}\left(\theta_{i,p}^{true} - \overline{\theta}_{i,p}^{post}\right)}$, $\theta_{i,p}^{true}$ is the true value of soil parameter θ_i for a given plot *p*, and $\overline{\theta}_{i,p}^{post}$ is the corresponding estimate given by the bayesian method. RE_i quantifies how much the estimate given by the bayesian method improves (RE_i<1) or not (RE_i≥1) the prior knowledge of the parameter value. For the synthetic database *P* is 50 and for the real database, it is 4, 10 and 12 for experiments on wheat, sugar beet and accumulated observations respectively.

2.3.4 Criterion expressing the quality of prediction

In a similar way we proposed a criterion, denoted REP_j (for Relative Error of Prediction) to quantify the quality of the prediction of the 3 agro-environmental variables defined above. It computes the ratio of the error of prediction obtained from the mean of the posterior distributions of the parameters, $\overline{\theta}^{post}$, to that obtained from the mean of the prior distributions, $\overline{\theta}^{prior}$:

$$REP_j = \frac{RMSEP_j\left(\overline{\theta}^{post}\right)}{RMSEP_j\left(\overline{\theta}^{prior}\right)} \tag{6}$$

where $RMSEP_j\left(\overline{\theta}^{post}\right) = \sqrt{\frac{1}{P \times Q_j} \sum_{p=1}^{P} \sum_{q=1}^{Q_j} \left(f_j^q(\theta_p^{true}) - f_j^q(\overline{\theta}_p^{post})\right)^2}$, θ_p^{true} is the vector of the true values of soil parameters θ for a given plot *p*, $\overline{\theta}_p^{post}$ is the corresponding estimate given by the Importance Sampling method, and $f_j^q\left(\theta_p^{true}\right)$ is assumed to be one of the Q_j observations of the predicted variable *j*, for the p^{th} vector of true values of soil parameters. For the synthetic database, Q_j is equal to 120 and for the real database its value depends on the number of available observations for the prediction as shown in Table 3.

3. Results and discussion

3.1 Synthetic experiments

3.1.1 Parameter estimates

The results of parameter estimation are shown in Figure 1 as the mean of the 4 weather patterns for one season of observation on wheat and sugar beet, 2 soil depths and 3 sets of observations. The results show that the observations provide substantial information for only 4 of the 7 soil parameters: *epc(2)*, *HCC(1)*, *HCC(2)* and *Hinit*. For these parameters the maximal gain from the prior information is between 38 and 70%. The other 3 parameters are difficult to estimate with these observations: for *argi*, *Norg* and *NO3init* the maximal gain is between 5 and 12%. These results can be explained by the different levels of sensitivities of the observable variables to these parameters as shown in (Varella et al., 2008).

The crop observed plays an important role in the quality of the parameter estimation process. Observations on sugar beet always provide better estimates of the parameters related to permanent properties (*epc(2)*, *HCC(1)* and *HCC(2)* and to a lesser extent *argi* and *Norg*), while observations on wheat always provide better estimates of the initial conditions (*Hinit* and to a lesser extent *NO3init*). These results can also be explained by the sensitivities of the observed variables to the parameters. As sugar beet is a summer crop (subjected to water stress), the soil water retention properties are more important for its growth than for winter wheat (usually not subjected to water stress in this region). The observed variables of sugar beet are thus more sensitive to the soil's permanent properties, and the corresponding observations contain more information to estimate the related parameters than those related to initial conditions. The lower sensitivity of sugar beet to initial conditions as compared to wheat is because the rainfall is often heavy in spring after sugar beet sowing, reducing the importance of initial conditions compared with wheat, which is sown in autumn when the weather is drier.

The soil depth only affects the parameter *epc(2)* whose quality of estimation is clearly better when the soil is shallow: for observations on wheat, the gain with respect to the prior information is around 25% for shallow soil and 5% for deep soil The corresponding values for sugar beet are 50% and 25% respectively. For the other parameters, the criterion ER_i calculated for the two soil depths has almost the same value.

Finally, the size of the observation set only slightly affects the quality of parameter estimation even though, as expected, the bigger the observation set the better is the parameter estimation. The gain is always limited to a few percent.

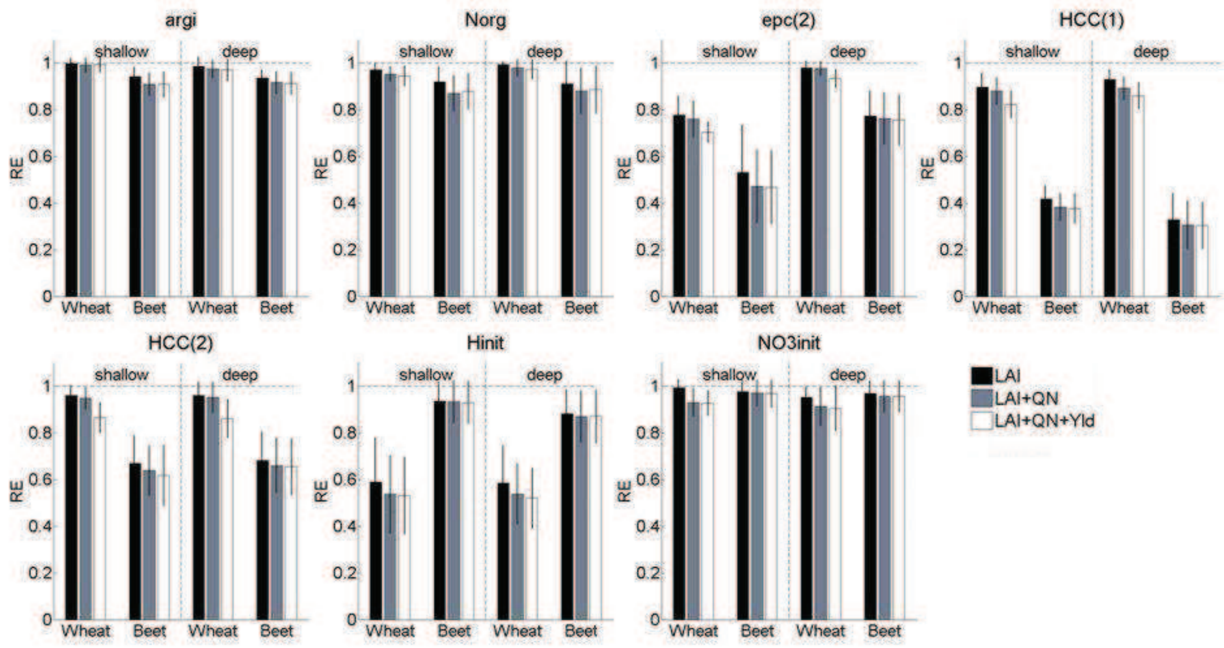

Figure 1. Results of the parameter estimation averaged over the 4 weather patterns for one season of observation on wheat (Wheat) and sugar beet (Beet), the 2 soil depths and the 3 sets of observations.

The effect of weather on the quality of parameter estimation is then analyzed for the mean of the 3 observation sets for one season of observation on wheat on a shallow soil, as shown in Figure 2. Whatever the weather pattern, the parameter set related to soil mineral nitrogen content is difficult to estimate (Fig. 2a): the lowest value of RE_i is over 0.9. The weather only affects the quality of estimation of the parameters related to soil water content: the drier the season, the better is the estimate of these parameters. The effect of weather on *Hinit* is striking: RE_i is about 0.77 on average for the wet weather pattern and about 0.4 for the dry weather pattern (which means a gain of 23% in the first case and 60% in the second). As the sensitivity of the observable variables to the parameters related to the water content is higher in dry conditions than in wet ones, the quality of estimates of these parameters is better in the first case.

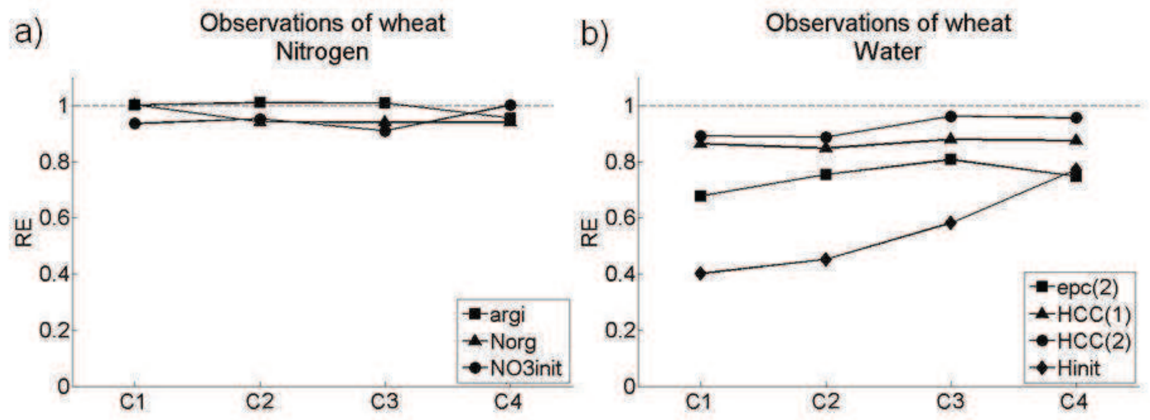

Figure 2. Effect of the weather on the parameter estimation averaged over the 3 observation sets, for one season of observation on wheat and a shallow soil. The parameters involved in the soil nitrogen content are in a) and those involved in the soil water content are in b). For the wheat, C1 characterizes a dry season, C2 a medium-dry season, C3 a medium-wet season and C4 a wet season.

3.1.2 Prediction of variables of interest

The quality of prediction will now be analyzed. Figure 3 shows the results in terms of REP_j for the prediction of the variables of interest for both crops, by using the estimated values of the permanent soil parameters (*argi*, *Norg*, *epc(2)*, *HCC(1)* and *HCC(2)*) instead of their prior values. The initial conditions (*Hinit* and *NO3init*), considered as unknown, are fixed at the mean of their distributions. As in Fig. 1, REP_j is calculated for one season of observations on wheat and sugar beet, 2 soil depths and 3 sets of observations. REP_j is then averaged over the 4 weather patterns. For both crops it can be seen that 2 of the 3 variables of interest (*Yld* and *Prot*) have a significantly lower REP_j and therefore a greatly improved quality of prediction when using the estimated values, as compared to when using prior information on the parameters. Note that the prediction quality of *Yld* is better when the wheat is predicted than when the sugar beet is predicted because the sugar beet crop is more sensitive to the initial conditions which are considered to be unknown in these results. The output variable *Nit* is only slightly affected, if at all, by the estimation of the soil parameters because it is sensitive to the initial conditions, which are not estimated, and not to the permanent parameters.

Through the estimation of the permanent soil properties, the size of the observation set slightly influences the quality of prediction. The biggest improvement between two observation sets is for parameters estimated from observations on wheat when the wheat growth is predicted for a shallow soil and the output variable is *Yld*: REP_j is about 0.66 for set #2 and about 0.52 for set #3.

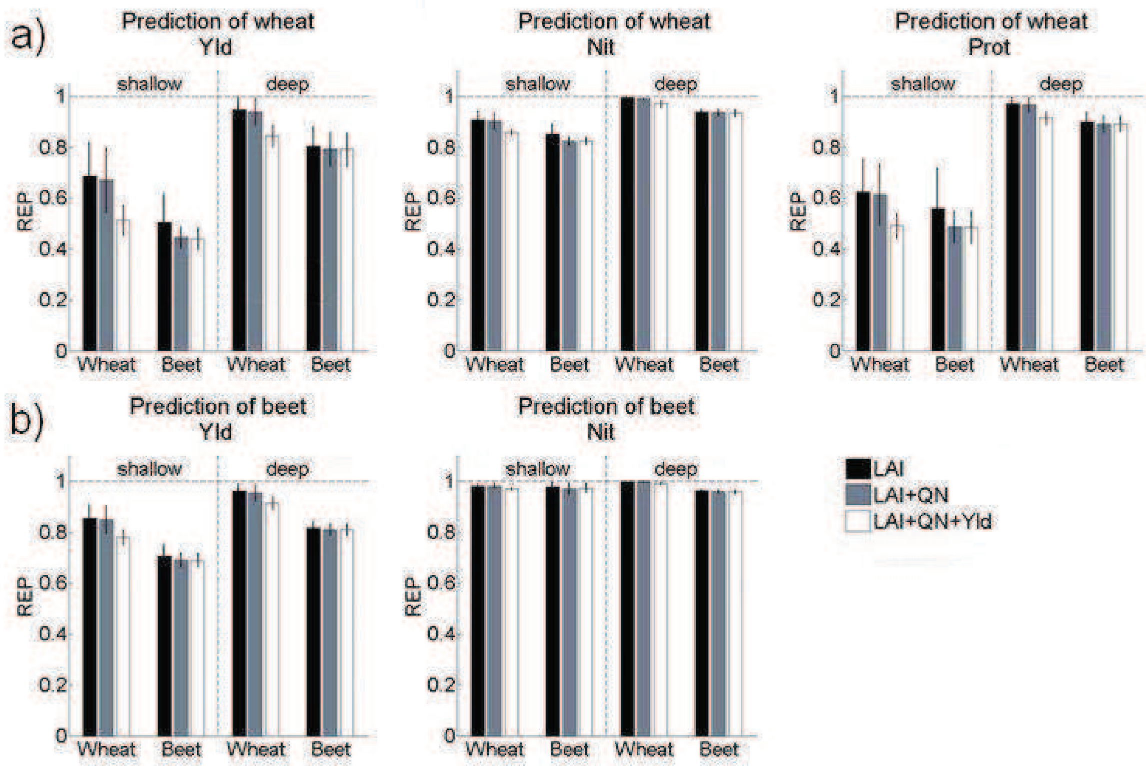

Figure 3. Results of the prediction of the variables of interest for a) wheat and b) sugar beet. The predictions are simulated by using the estimated values of the permanent soil parameters and by fixing the initial conditions at the mean of their distributions. The results are averaged over the 4 weather patterns for one season of observation on wheat (Wheat) and sugar beet (Beet), the 2 soil depths and the 3 sets of observations.

The crop on which observations are made for parameter estimation also plays an important role in the quality of prediction. Because observations on sugar beet improve the quality of permanent parameter estimates, the results show that the predictions are better when the soil parameters are estimated with observations made on sugar beet. This effect can be seen on the 3 predicted variables but especially on *Yld*. It is particularly striking for the prediction of *Yld* for wheat for observation set #2: the gain from prior information varies from 33 to 55% when using observations on sugar beet rather than wheat. It is striking that the improvement in the wheat prediction is greater by using observations on sugar beet than on wheat. As the sensitivity of the observed variables on sugar beet to the permanent parameters is higher than the observed variables on wheat, the estimation of these parameters is thus better, implying a smaller uncertainty on the prediction of output variables in this case.

The soil depth affects the quality of the prediction, especially for the output variables *Yld* and *Prot*, which have a lower REP_j when the soil is shallow. It is not surprising, accordingly to the results of parameter estimation, because the parameter *epc(2)* gives better estimates in this case and because these outputs are quite sensitive to this parameter. The output variable *Nit* is not affected by the soil depth because of its lack of sensitivity to *epc(2)*.

The effect of weather on the quality of prediction of the variables of interest is illustrated for wheat in Figure 4 for a shallow soil and one year of observations: the results are the means of the 3 observation sets. The drier the weather (from wet C4 to dry C1) the better is the prediction, which follows from the previous results showing that the quality of estimate of permanent parameters related to the water content is affected by the weather. The REP_j value of the output variables *Yld* and *Prot* decreases significantly from 0.71 to 0.5 and 0.65 to 0.47 respectively. As before, the output variable *Nit* is not accurately predicted and is only slightly affected by the weather: REP_j decreases from 0.92 to 0.86.

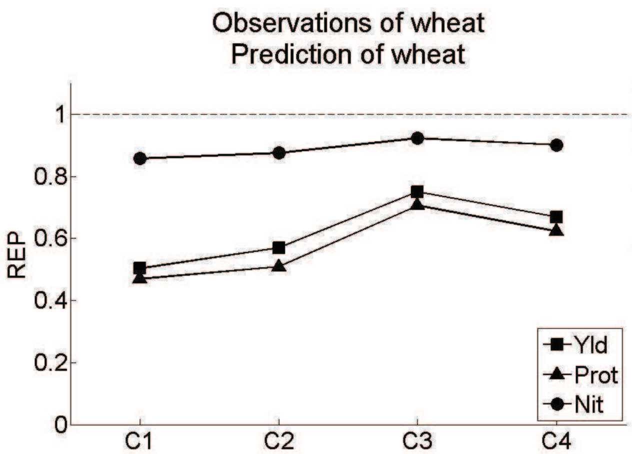

Figure 4. Effect of the weather on prediction of the variables of interest of the wheat crop when the initial conditions are unknown. The results are averaged over the 3 observation sets for one observed season on wheat crop and a shallow soil. For the wheat, C1 characterizes a dry season, C2 a medium-dry season, C3 a medium-wet season and C4 a wet season.

Given that one season of observations on sugar beet allows better prediction (see Fig. 3), we investigated the marginal improvement due to adding another observation season on sugar beet or wheat. Figure 5a and 5c shows the results in terms of REP_j for the prediction of the variables of interest for both crops for a shallow soil and the observation set #3 averaged over the different weather patterns. In this case, as in the previous one, the initial conditions are assumed to be unknown. The results show that the addition of another observation season either on sugar beet or on wheat does not significantly improve the prediction obtained using only one observation season on sugar beet. They confirm that the better improvement in prediction is made for the wheat crop and that the output variable *Nit* is difficult to predict correctly.

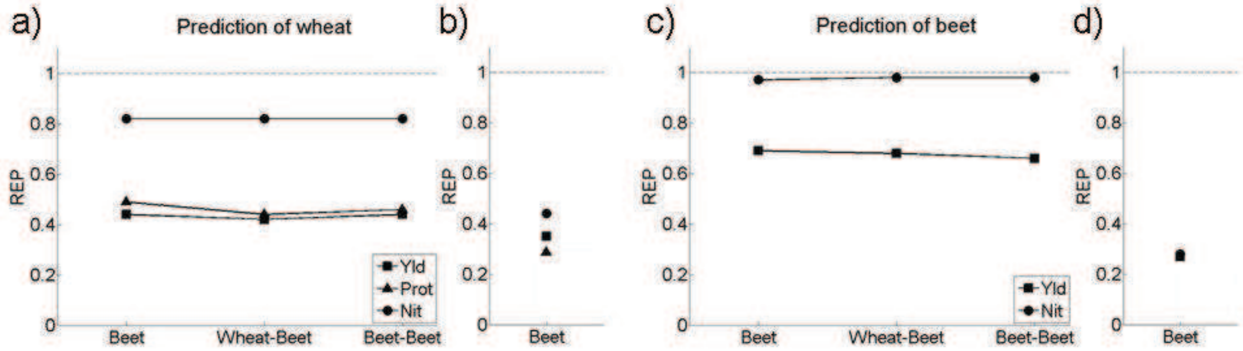

Figure 5. Results of the prediction of the variables of interest for the wheat crop when the initial conditions are a) unknown b) known, and for the sugar beet crop when the initial conditions are c) unknown d) known. One season of observations on sugar beet (Beet) and two pairs of combined observation seasons are considered: wheat for the first season and sugar beet for the second season (Wheat-Beet), and sugar beet for both seasons (Beet-Beet). The results are averaged over the various weather patterns and shown for a shallow soil and the third set of observations.

For crop monitoring applications, the initial conditions may be considered as known (from measurements). Fig. 5b and 5d show the results obtained for the same situation as before, with the initial conditions set to their true values. In this case, the reduction of the error of prediction is striking for the sugar beet crop, which is more sensitive to the initial conditions than wheat. All the output variables are affected but the best improvement is for *Nit* which is the most sensitive to the initial conditions: REP_j decreases from 0.82 to 0.44 when wheat is predicted and from 0.97 to 0.28 when sugar beet is predicted.

3.2 Real experiments

3.2.1 Parameter estimates

The results of parameter estimation on the real experiments at Bruyères are presented in Figure 6 for different observation sets. In every case *HCC(1)* is the best estimated parameter (RE_i between 0.85 and 0.54). The parameters involved in the soil mineral nitrogen content, *argi* and *Norg*, can be slightly estimated (gain between 2 and 22%), and the parameter *epc(2)* can be estimated only when observations on sugar beet crop are used (RE_i between 0.7 and 0.62 instead of 1.15 when observations on wheat are used). The other parameters, *HCC(2)*, *Hinit* and *NO3init*, can not be estimated and have a RE_i value of 1 or slightly more. This is mainly due to a lack of observations early in wheat growth, which is when the observed variables are sensitive to *Hinit* and *NO3init*. As compared to the results obtained on synthetic data, we would argue that several types of bias have to be considered in addition on these results : the bias due to the STICS model error on the output simulation, the omission bias(Miller, 2002) and the bias on the observations. The addition of such biases increases the error of the estimated values of the parameters. This is clearly visible on the parameters with an RE_i greater than 1 (*epc(2)* and *HCC(2)* for the wheat observation set and the initial conditions *Hinit* and *NO3init*): in these cases these biases do not allow to have a $RMSE\left(\overline{\theta}_i^{post}\right)$ lower than $RMSE\left(\overline{\theta}_i^{prior}\right)$.

Observations on sugar beet always lead to better permanent parameter estimates than observations on wheat. The effect of the crop observed varies greatly according to the parameter: RE_i from 0.88 to 0.78 for *Norg*, from 1.15 to 0.62 for *epc(2)*, from 2.2 to 0.95 for *HCC(2)*. For the initial conditions, the lack of observations in the early period of the growing season prevents us from assessing the effect of the crop observed. Accumulating observations on wheat and sugar beet crops is always preferable to using only wheat observations, and the gain obtained varies according to the parameters. Nevertheless, observations on a single sugar beet crop season provide more information for estimating the permanent parameters than the accumulation of observations on wheat and sugar beet crops. The gain from using prior information is 5% for *HCC(2),* 20% for *argi*, 22% for *Norg*, 38% for *epc(2)* and 46% for *HCC(1)*.

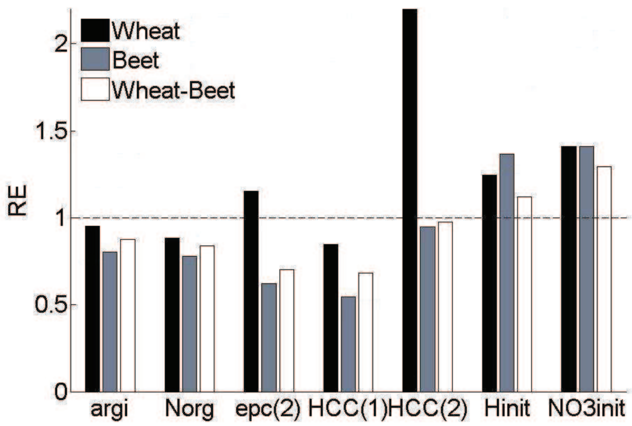

Figure 6. Results of parameter estimation on the real experiments of Bruyères, for one season of observation on wheat (Wheat) or sugar beet (Beet) and for accumulated observations on wheat and sugar beet crops (Wheat-Beet).

Some of these results are quite consistent with those obtained from experiments on the same configurations of synthetic data, i.e. wet weather patterns and deep soils, as in the real data. Firstly, for a given observation set, *HCC(1)* is the best estimated permanent soil parameter (see Fig. 1 and 6) and can thus be estimated with great confidence. Secondly, observations on sugar beet provide very informative data for estimating the permanent soil parameters and allow better estimates of those parameters than observations on wheat. For both synthetic and real data, *argi* and *Norg* can be slightly estimated**;** the observed crop has little influence on their estimates. Except for observations on wheat, the results of the estimation of *epc(2)* and *HCC(2)* are consistent with those obtained on the synthetic data: *epc(2)* is well estimated while *HCC(2)* is quite difficult to predict.

Nevertheless, there are some differences from the results obtained on the synthetic data. For observations on wheat, *epc(2)* and *HCC(2)* have a very high value (above 1) of ER_i for real data while they are reasonably well estimated (about 0.75 for *epc(2)* and 0.9 for *HCC(2)*) for synthetic data (see Fig. 6). This could be explained by the presence of increased bias (model error bias, omission bias or observation bias) in estimating these parameters. The last difference concerns the initial conditions *Hinit* and *NO3init* and the impossibility of estimate these parameters for real data

(ER_i above 1) due to the lack of early observations, unlike the case of the synthetic data.

3.2.2 Prediction of variables of interest

Figure 7 compares observed variables with those simulated by STICS using prior information (Fig. 7a, 7b, 7c) and estimated values (Fig. 7d, 7e, 7f) of the soil parameters. The initial conditions are assumed to be unknown. For the simulations using prior information, the results are quite good for wheat *Yld* and *Prot* (Fig. 7a, 7c) although *Yld* is slightly under-estimated. The simulations are improved by the parameter estimation process (Fig. 7d, 7f) and *Yld* is no longer under-estimated: the model error is low and the simulation error is mainly due to inaccurate soil parameter values, even though the prior information is fairly suitable. Conversely, sugar beet *Yld* is less accurately simulated by the model and is not improved by parameter estimation because the model error is higher than for wheat *Yld* and cannot be reduced by estimating the soil parameters. These results are partly due to a better previous calibration of the plant parameters for wheat than for sugar beet (see Section 2.1.1), involving a smaller model error for wheat crop simulation. *Nit* (Fig 7b) is poorly predicted by the model for both wheat and sugar beet and is not improved by parameter estimation (Fig 7e): this confirms the inability of STICS to predict this variable satisfactorily (Houlès et al., 2004).

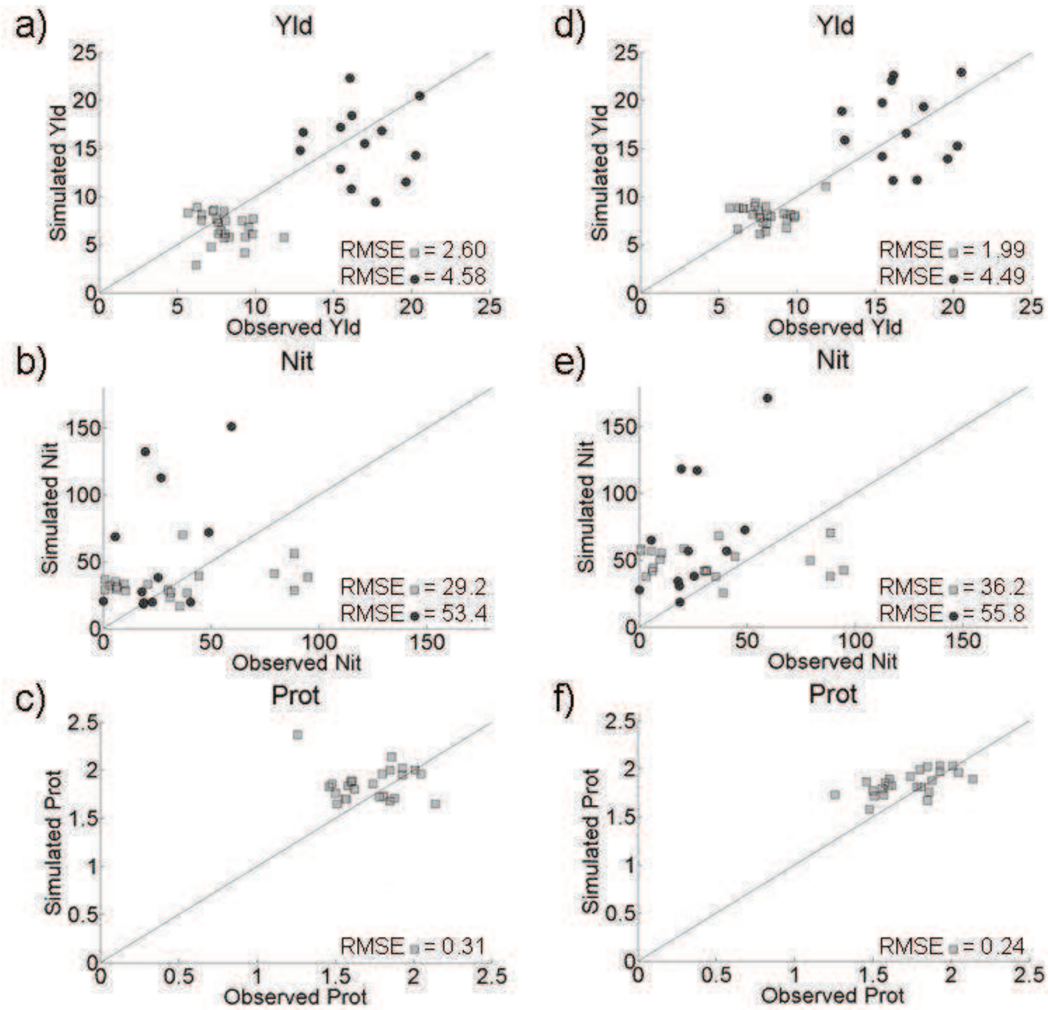

Figure 7. Comparison of observed and simulated variables *Yld*, *Nit* and *Prot*, obtained with the STICS model on the Bruyères database. a), b) and c) are for simulations using the prior values of the soil parameters and d), e) and f) are for those using the estimated values (those estimated with observations accumulated on wheat and sugar beet crop). In both cases, the initial conditions are unknown. □ = wheat, ○ = sugar beet.

An averaged representation of these results is given in Figure 8 where the prediction of the variables of interest for both crops after estimating the soil permanent parameters (*argi*, *Norg*, *epc(2)*, *HCC(1)* and *HCC(2)*) is shown. In a first attempt, the initial conditions (*Hinit* and *NO3init*) are assumed to be unknown and set to their prior values (Fig. 8a and 8b). The prediction of wheat variables *Yld* and *Prot* is generally improved (by about 23%) by estimating the permanent soil parameters as compared to the results obtained with prior information: this is due to their sensitivity to these parameters. For sugar beet the improvement is not significant for the output variable *Yld* (the lower REP_j value is about 0.92). However the prediction of the output *Nit* can never be improved because of a lack of sensitivity to the permanent soil parameters but also because of a large STICS model error involved in its simulation, as stated earlier.

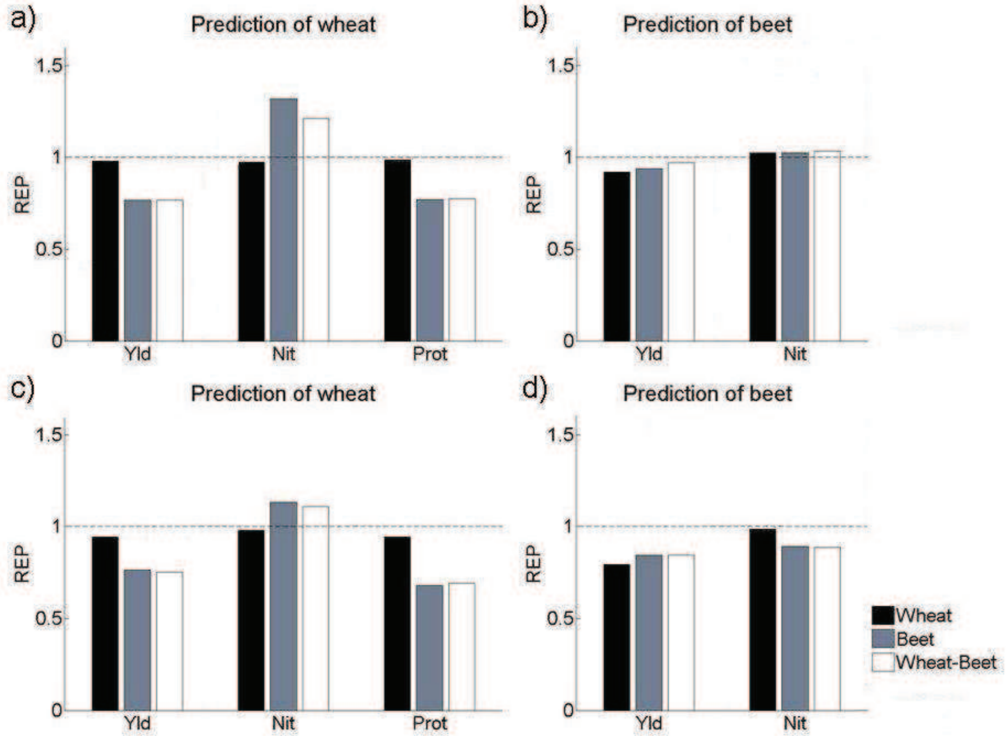

Figure 8. Results of the prediction of the variables *Yld*, *Nit* and *Prot*, on the Bruyères database, for wheat (a and c) and sugar beet (b and d), when the initial conditions are unknown (a and b) or known (c and d). The results are shown for one season of observation on wheat (Wheat) or sugar beet (Beet) and for accumulations of observations on wheat and sugar beet crops (Wheat-Beet).

The effect of the crop observed on the prediction of wheat growth is striking, but not on that of sugar beet. The error of the prediction of wheat *Yld* and *Prot* decreases considerably when using observations on sugar beet rather than on wheat: it falls from 0.98 to 0.77 for *Yld* and from 0.99 to 0.77 for *Prot*. However the observed crop does not affect the prediction of sugar beet growth or its error. The accumulation of observations on wheat and sugar beet crops improves the prediction of wheat *Yld* and *Prot* as compared to the use of a single observation season on wheat, while it has no effect on the prediction of sugar beet growth. Finally, accumulating observations on different crops does not improve the results obtained with sugar beet observations alone.

In a second step the initial conditions were known and fixed at their actual measured values (Fig. 8c and 8d). This led to an all-round improvement. The output

Nit for both crops is the most affected variable because of its great sensitivity to the initial conditions. The prediction error of sugar beet *Nit* can be significantly reduced when observations on sugar beet are considered and when the initial conditions are known (REP_j is about 0.78). Moreover, the *Yld* output is less improved for wheat than for sugar beet for which the error is reduced from an REP_j of about 0.92 when the initial conditions are unknown to about 0.79-0.85 when they are known. As stated before, sugar beet *Yld* is more sensitive to the initial conditions

Some of these results are quite consistent with those obtained on the synthetic data for wet conditions and deep soil. First, both experiments show that observations on sugar beet give a better prediction of wheat *Yld* and *Prot* (see Fig. 3 and 7) than those on wheat. Moreover, both studies show that combining or accumulating observations over several wheat and sugar beet crop seasons does not improve predictions any more than considering observations from a single sugar beet crop season, revealing the importance of this kind of observation for the parameter estimation and the consequent prediction. The output variable *Nit* proves to be very difficult to predict using these kinds of observations and the STICS model. The last main consistent result concerns the effect of the knowledge of the initial conditions on the reduction of the prediction error. This knowledge can reduce the prediction error of all the output variables but particularly for sugar beet because they are very sensitive to the initial conditions for this crop. The main difference between the results obtained on the real and synthetic data concerns the prediction of sugar beet growth: there is no improvement in the real case for *Yld* with sugar beet observations. In this case, a high STICS model error for the simulation of sugar beet *Yld* (see results presented above for Fig. 7) can biases these results, which is not the case for the synthetic experiments.

4. Conclusions

The prediction of agro-environmental variables can be improved by estimating the soil parameters from observations on crops and the STICS model. The main results found by using the synthetic observations are confirmed by the results found by using real observations, revealing the value and complementary nature of the two studies. The synthetic study enables the effects of a large number of possible observation sets to be explored, while the real study can validate the effects of a restricted number of observation sets, taking account of the model error. The estimation of the parameters related to the soil water content can be greatly improved and the prediction of the variables of interest is thus also affected because of their sensitivity to these parameters, particularly for yield and grain protein. As Braga (2004) and Varella (2008) showed in their studies, the quantity of information available on the parameters is different according to the type of the observation set. In this study, we showed that observation sets obtained on sugar beet, in a dry season or on a shallow soil, allow better parameter estimation and thus better predictions. For even drier weather conditions or shallower soils, better estimation and prediction could be expected.

The results of parameter estimation and variable prediction are closely linked to some hypotheses made in this paper. The crucial hypothesis concerns the knowledge of the initial conditions (*Hinit* and *NO3init*) and affects only the prediction quality. If the initial conditions are assumed to be unknown, the prediction error takes into account the error involved by fixing them at a nominal value (Sobol et al., 2007). The error can be thus reduced if the initial conditions values are determined by the measurements of water and mineral nitrogen content at the beginning of the crop simulation. When a given output variable is sensitive to the initial conditions, such as *Nit* or sugar beet *Yld*, its prediction error is even more reduced by making use of this information. For crop monitoring applications it is thus important to deal with practical and financial constraints for measuring the initial conditions.

The results of parameter estimation and variable prediction are also closely linked to the prior information used to estimate parameters. In this study, this information came from examining the distribution of the soil measurements made on a soil similar to that of the real database considered here. However, if time and financial constraints allow it, the prior information could also be provided by making a partial soil analysis at certain locations within the study area of the real database and using it to provide a mean and variance for soil parameter values which might lead to better parameter estimates. The results of the parameter estimation and the prediction could be even more improved if new information could be acquired on some parameters correlated with others. This may be done by measuring them and fixing them at the measured values, improving the prior information, or adding new observations related to these parameters in the estimation process (Varella et al., 2008). The uncertainty of parameter estimation should be thus reduced and the performance of the prediction should be greatly improved.

From a practical point of view, the study shows how the observation set chosen can influence parameter estimation and prediction. If the user is able to characterize the observation set in terms of the crop observed, the weather pattern and soil depth, it should be possible to determine which parameters can be estimated and which variables of interest can be predicted with a given accuracy. Labor and costs for collecting the data could be minimized and the performance of the methodology could be maximized: the user should seek to collect the data when the soil and weather conditions are optimal, such as when a sugar beet crop is grown in a dry season.

Acknowledgements: The financial support provided by CNES and Arvalis-Institut du Végétal is gratefully acknowledged.

References

Baret, F., Houles, V., Guerif, M., 2006. Quantification of plant stress using remote sensing observations and crop models: the case of nitrogen management. Symposium on Imaging Techniques for Understanding Plant Reponses to Stress held at the Society-for-Experimentaal-Biology Meeting, Canterbury, ENGLAND, pp. 869-880.

Beaudoin, N., Launay, M., Sauboua, E., Ponsardin, G., Mary, B., 2008. Evaluation of the soil crop model STICS over 8 years against the "on farm" database of Bruyeres catchment. European Journal of Agronomy 29, 46-57.

Beaudoin, N., Saad, J.K., Van Laethem, C., Machet, J.M., Maucorps, J., Mary, B., 2005. Nitrate leaching in intensive agriculture in Northern France: Effect of farming practices, soils and crop rotations. Agriculture, Ecosystems & Environment 111, 292-310.

Beven, K., Binley, A., 1992. The Future of Distributed Models - Model Calibration and Uncertainty Prediction. Hydrological Processes 6, 279-298.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology 249, 11-29.

Blackmore, S., Moore, M., 1999. Remedial correction of yield map data. Precision Agriculture 1, 53-56.

Braga, R.P., Jones, J.W., 2004. Using Optimization to Estimate Soil Inputs of Crop Models for Use in Site-Specific Management. Transactions of the ASAE 47, 1821-1831.

Brisson, N., Ruget, F., Gate, P., Lorgeou, J., Nicoullaud, B., Tayot, X., Plenet, D., Jeuffroy, M.H., Bouthier, A., Ripoche, D., Mary, B., Juste, E., 2002. STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. Agronomie 22, 69-92.

Calmon, M.A., Jones, J.W., Shinde, D., Specht, J.E., 1999. Estimating parameters for soil water balance model using adaptative simulated annealing. Applied Engineering in Agriculture 15, 703-7013.

Ferreyra, R.A., Jones, J.W., Graham, W.D., 2006. Parameterizing Spatial Crop Models with Inverse Modeling: Sources of Error and Unexpected Results. Transactions of the ASABE 49, 1547-1561.

Flenet, F., Villon, P., Ruget, F.O., 2003. Methodology of adaptation of the STICS model to a new crop: spring linseed (Linum usitatissimum, L.). STICS Workshop, Camargue, FRANCE, pp. 367-381.

Gabrielle, B., Roche, R., Angas, P., Cantero-Martinez, C., Cosentino, L., Mantineo, M., Langensiepen, M., Henault, C., Laville, P., Nicoullaud, B., Gosse, G., 2002. A priori parameterisation of the CERES soil-crop models and tests against several European data sets. Agronomie 22, 119-132.

Gaucherel, C., Campillo, F., Misson, L., Guiot, J., Boreux, J.-J., 2008. Parameterization of a process-based tree-growth model: Comparison of optimization, MCMC and Particle Filering algorithms. Environmental Modelling & Software 23, 1280-1288.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1995. Markov Chain Monte Carlo in Practice. Chapman & Hall, London.

Golovko, L., Pozdnyakov, A.I., 2007. Electrical geophysical methods in agriculture. Progress of Information Technology in Agriculture, 457-471.

Gras, R., Erhwein, J., Hermant, G., Hebert, J., 1961. Carte des sols du département de l'Aisne. Laon n 5-6. Chambre d'agriculture de l'Aine.
Guérif, M., Beaudoin, N., Durr, C., Machet, J.M., Mary, B., Michot, D., Moulin, D., Nicoullaud, B., Richard, G., 2001. Designing a field experiment for assessing soil and crop spatial variability and defining site specific management strategies. Proceedings 3rd European Conference on Precision Agriculture, Montpellier, France, pp. 677-682.
Guérif, M., Houlès, V., Makowski, D., Lauvernet, C., 2006. Data assimilation and parameter estimation for precision agriculture using the crop model STICS. In: Wallach, D., Makowski, D., Jones, J.W. (Eds.), Working with Dynamic Crop Models. Elsevier, pp. 395-401.
Hadria, R., Khabba, S., Lahrouni, A., Duchemin, B., Chehbouni, A., Carriou, J., Ouzine, L., 2007. Calibration and validation of the STICS crop model for managing wheat irrigation in the semi-arid Marrakech/Al Haouzi plain. Arabian Journal for Science and Engineering 32, 87-101.
Houborg, R., Boegh, E., 2008. Mapping leaf chlorophyll and leaf area index using inverse and forward canopy reflectance modeling and SPOT reflectance data. Remote Sensing of Environment 112, 186-202.
Houlès, V., Mary, B., Guérif, M., Makowski, D., Juste, E., 2004. Evaluation of the crop model STICS to recommend nitrogen fertilization rates according to agro-environmental criteria. Agronomie 24, 1-9.
Irmak, A., Jones, J.W., Batchelor, W.D., Paz, J.O., 2001. Estimating Spatially Variable Soil Properties for Application of Crop Models in Precision Farming. Transactions of the ASAE 44, 1343-1353.
King, D., Daroussin, J., Tavernier, R., 1994. Development of a Soil Geographic Database from the Soil Map of the European Communities. Catena 21, 37-56.
Lagacherie, P., Baret, F., Feret, J.B., Netto, J.M., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. Remote Sensing of Environment 112, 825-835.
Launay, M., Guérif, M., 2003. Ability for a model to predict crop production variability at the regional scale: an evaluation for sugar beet. Agronomie 23, 135-146.
Machet, J.M., Couturier, A., Beaudoin, N., 2007. Cartographie du rendement du blé et des caractéristiques qualitatives des grains. In: Guérif, M., King, D. (Eds.), Agriculture de précision. Quae, Versailles, pp. 131-144.

Makowski, D., Naud, C., Jeuffroy, M.H., Barbottin, A., Monod, H., 2006. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. Reliability Engineering & System Safety 91, 1142-1147.

Makowski, D., Wallach, D., Tremblay, M., 2002. Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods. Agronomie 22, 191-203.

Miller, A., 2002. Subset Selection in Regression. Chapman & Hall.

Moulin, S., Zurita, R.M., Guérif, M., 2007. Estimation de variables biophysiques du couvert par ajustement de modèles de transfert radiatif sur des réflectances. In: Guérif, M., King, D. (Eds.), Agriculture de précision. Quae, Versailles, pp. 97-113.

Murphy, B., Geeves, G., Miller, M., Summerell, G., Southwell, P., Rankin, M., 2003. The application of pedotransfer functions with existing soil maps to predict soil hydraulic properties for catchment-scale hydrologic and salinity modelling. International Congress on Modelling and Simulation, Townsville, AUSTRALIA, pp. 502-507.

Reynolds, C.A., Jackson, T.J., Rawls, W.J., 2000. Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. Water Resources Research 36, 3653-3662.

Ruget, F., Brisson, N., Delecolle, R., Faivre, R., 2002. Sensitivity analysis of a crop simulation model, STICS, in order to choose the main parameters to be estimated. Agronomie 22, 133-158.

Sobol, I.M., Tarantola, S., Gatelli, D., Kucherenko, S.S., Mauntz, W., 2007. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. Reliability Engineering & System Safety 92, 957-960.

Timlin, D., Pachepsky, Y., Walthall, C., Loechel, S., 2001. The use of a water budget model and yield maps to characterize water availability in a landscape. Soil & Tillage Research 58, 219-231.

Tremblay, M., Wallach, D., 2004. Comparison of parameter estimation methods for crop models. Agronomie 24, 351-365.

Varella, H., Guérif, M., Buis, S., 2008. Global Sensitivity Analysis (GSA) measures the quality of parameter estimation. Case of soil parameter estimation with a crop model., EGU, Vienna, Austria.
Weiss, M., Baret, F., 1999. Evaluation of canopy biophysical variable retrieval performances from the accumulation of large swath satellite data. Remote Sensing of Environment 70, 293-306.