

Simulation de la croissance du blé à l'aide de modèles écophysologiques : synthèse bibliographique des méthodes, potentialités et limitations

Benjamin Dumont ⁽¹⁾, Françoise Vancutsem ⁽²⁾, Benoit Seutin ⁽²⁾, Bernard Bodson ⁽²⁾, Jean-Pierre Destain ⁽³⁾, Marie-France Destain ⁽¹⁾

⁽¹⁾ Univ. Liège - Gembloux Agro-Bio Tech. Unité de Mécanique et Construction. Passage des Déportés, 2. B-5030 Gembloux (Belgique). E-mail : benjamin.dumont@ulg.ac.be

⁽²⁾ Univ. Liège - Gembloux Agro-Bio Tech. Unité de Phytotechnie des Régions tempérées. Passage des Déportés, 2. B-5030 Gembloux (Belgique).

⁽³⁾ Centre wallon de Recherches agronomiques. Département Agriculture et Milieu naturel. Rue de Liroux, 9. B-5030 Gembloux (Belgique).

Reçu le 23 septembre 2011, accepté le 3 juillet 2012.

Les modèles de culture ou modèles écophysologiques décrivent la croissance et le développement de cultures en interaction avec leurs conditions agro-environnementales (sol, climat et environnement proche de la plante). Étant donné le grand nombre de variables d'entrée et de paramètres requis, la mise en œuvre de ces modèles peut s'avérer délicate et il arrive fréquemment que des écarts importants existent entre les valeurs mesurées et simulées. Cet article vise à mettre en évidence les différentes sources d'incertitude liées au fonctionnement des modèles de culture ainsi que les méthodes existantes qui permettent de pallier ou, à tout le moins, de prendre en compte ces sources d'erreurs. L'article présente d'abord une revue bibliographique synthétisant la structure mathématique générale des modèles de cultures. Les principaux critères permettant de juger la qualité des modèles en phase de calibration ou de validation sont ensuite exposés. Enfin, différentes méthodes permettant d'améliorer les modèles sont abordées. On distingue d'une part les méthodes d'estimation paramétrique fréquentielles et bayésiennes et d'autre part, les techniques d'assimilation de données.

Mots-clés. *Triticum*, blé, plante de culture, sol, croissance, calibrage, méthode d'optimisation.

Wheat growth simulation using crop models: A review of methods, and their potential and limitations. Crop models describe the growth and development of a crop interacting with its surrounding agro-environmental conditions (soil, climate and the close conditions of the plant). However, the implementation of such models remains difficult because of the high number of explanatory variables and parameters. It often happens that important discrepancies appear between measured and simulated values. This article aims to highlight the different sources of uncertainty related to the use of crop models, as well as the actual methods that allow a compensation for or, at least, a consideration of these sources of error during analysis of the model results. This article presents a literature review, which firstly synthesises the general mathematical structure of crop models. The main criteria for evaluating crop models are then described. Finally, several methods used for improving models are given. Parameter estimation methods, including frequentist and Bayesian approaches, are presented and data assimilation methods are reviewed.

Keywords. *Triticum*, wheats, crops, soil, growth, calibration, optimization methods.

1. INTRODUCTION

Depuis plus de 25 ans, des modèles de culture, aussi appelés modèles écophysologiques, fournissent une représentation conceptuelle du système sol – culture dans différentes conditions pédo-climatiques. On distingue des modèles spécifiques dédiés à l'étude de la croissance du blé ou du maïs, comme CERES-

Wheat (Ritchie et al., 1984 ; Singh et al., 2008), ARCWHEAT (Weir et al., 1984), CERES-Maize (Jones et al., 1986) et des modèles plus génériques tels que EPIC (William et al., 1989), WOFOST (Van Diepen et al., 1989), DAISY (Hansen et al., 1990), STICS (Brisson et al., 1998 ; Brisson et al., 2003 ; Brisson et al., 2008) et SALUS (Basso et al., 2005 ; Basso et al., 2009).

Traditionnellement, les modèles écophysiologiques jouent un double rôle. Ils constituent une approche privilégiée pour mettre au point les systèmes de culture en présence de combinaisons complexes de techniques et de successions de cultures, face à la variabilité du sol et du climat (actuel et futur). Outre le rendement, ils permettent de modéliser des critères de qualité comme la teneur en protéines des grains de blé qui est directement liée au fonctionnement carboné et azoté de la culture (Génard et al., 2006). D'autre part, ils sont capables de fournir des indicateurs d'impacts environnementaux tels que la lixiviation des nitrates, la salinisation des sols ou encore la séquestration du carbone dans le sol (Loyce et al., 2006 ; Beaudoin et al., 2008). Étant donné les exigences de la Directive européenne 91/676, les indicateurs relatifs à l'azote sont apparus les premiers dans les modèles. La quantité d'azote non prélevée par les racines au cours du cycle cultural et l'azote minéral contenu dans le sol après la récolte sont des variables privilégiées. Plus récemment, des adaptations ont été apportées aux modèles afin qu'ils fournissent une aide à la décision en temps réel, prenant en compte à la fois des aspects agronomiques et environnementaux. Les applications concernent par exemple la recommandation de doses de fertilisants adaptées à des critères agro-environnementaux (Houlès et al., 2004), la simulation de l'état de prairies pour anticiper la prise de décision en présence de fortes sécheresses (Ruget et al., 2009), etc.

Les modèles de culture peuvent donc se révéler être des outils précieux pour les acteurs de terrain. Cependant, ils sont constitués d'une suite d'équations imbriquées les unes dans les autres et leur utilisation peut se révéler très complexe. À travers une littérature abondante, le présent article vise à synthétiser, mettre en évidence et quantifier les limitations et les sources d'incertitude liées au fonctionnement de ces modèles. Dans un second temps sont présentées différentes techniques permettant d'améliorer la qualité prédictive de ces modèles.

2. LA STRUCTURE MATHÉMATIQUE DES MODÈLES ÉCOPHYSIOLOGIQUES

Les modèles de culture décrivent la croissance et le développement d'un système constitué d'une culture en interaction avec le sol et l'atmosphère. Ces modèles dynamiques continus sont constitués d'un ensemble d'équations différentielles du premier ordre, linéaires ou non, que l'on peut écrire sous la forme :

$$d\mathbf{U}(t) = g[\mathbf{U}(t), \mathbf{X}(t), \boldsymbol{\theta}] \cdot dt \quad (1)$$

où t est le temps, $\mathbf{U}(t)$ est le vecteur des variables d'état, $\mathbf{X}(t)$ est le vecteur des variables d'entrée ou variables

explicatives, $\boldsymbol{\theta}$ est le vecteur des paramètres et g est une fonction. Pour procéder à la simulation numérique, des équations aux différences sont utilisées et la forme générale d'un modèle de culture est alors la suivante :

$$\mathbf{U}(t + \Delta t) = \mathbf{U}(t) + g[\mathbf{U}(t), \mathbf{X}(t), \boldsymbol{\theta}] \cdot \Delta t \quad (2)$$

Δt est l'incrément de temps. Le plus souvent dans les modèles de culture, Δt est pris égal à un jour.

Les variables d'état $\mathbf{U}(t)$ sont des variables indépendantes qui décrivent l'état du système ou d'une partie de celui-ci (« sous-ensemble ») à un moment donné. Une variable d'état qui représente la sortie d'un sous-ensemble peut servir de variable d'entrée à un autre sous-ensemble. Ainsi, le LAI (*Leaf Area Index* ou indice de développement foliaire) est à la fois variable de sortie du sous-ensemble qui contrôle la croissance de la végétation et variable d'entrée du sous-ensemble qui contrôle l'accroissement de biomasse. Typiquement dans les modèles de culture, la biomasse totale et son allocation entre les différents organes végétaux (feuilles, tiges, racines et graines), le LAI, les stades de développement constituent des variables d'état. Les teneurs en eau des différents horizons de sol explorés par les racines en constituent d'autres, comme la teneur en azote minéral et en matière organique du sol. Notons qu'une variable d'état apparaît dans les deux membres des équations 2. Dans le membre de droite, elle représente une valeur à l'instant t , qui servira de base à l'estimation de la valeur à l'instant $t + \Delta t$ dans le membre de gauche.

Les variables d'entrée $\mathbf{X}(t)$ sont mesurées ou observées dans toutes les situations dans lesquelles le modèle est appliqué. Elles incluent les conditions initiales de l'expérimentation (la date de semis, etc.), les caractéristiques du sol (la teneur en eau journalière des différents horizons de la matrice sol, la réserve en eau utilisable par la plante, etc.), les variables climatiques (les températures journalières, etc.) et les variables relatives à la gestion de l'itinéraire technique (les dates et doses d'apport de fertilisant, etc.). Les variables climatiques sont encore appelées variables de conduite du modèle ou variables de forçage.

Les paramètres $\boldsymbol{\theta}$ se différencient des variables explicatives, en ce sens qu'ils sont constants pour toutes les situations d'intérêt étudiées. Ainsi, à titre d'exemple, les teneurs en eau à la capacité au champ et au point de flétrissement sont des constantes pour chaque horizon de sol. Ce sont les deux paramètres, à l'image de deux bornes, qui permettent de calculer la réserve en eau du sol disponible pour la plante.

Un modèle comportant trois variables d'état et destiné à estimer la production de biomasse et le LAI de manière très simplifiée est donné par Wallach et al. (2006), à titre d'exemple :

$$\begin{cases} B(j+1) = B(j) + \Delta B(j) \\ LAI(j+1) = LAI(j) + \Delta LAI(j) \\ TT(j+1) = TT(j) + \Delta TT(j) \end{cases} \quad (3)$$

avec B la biomasse, LAI l'indice foliaire (*Leaf Area Index*) et TT la somme des degrés-jours. Dans cet exemple, j est l'index correspondant au jour, le pas de temps Δt est égal à un jour. Les variations de degrés-jours $\Delta TT(j)$, de biomasse $\Delta B(j)$ et d'indice foliaire $\Delta LAI(j)$ en fonction des conditions environnementales et des caractéristiques des végétaux sont données par les équations suivantes :

$$\Delta TT(j) = \max \left[\frac{TMIN(j) + TMAX(j)}{2} - T_{base}, 0 \right] \quad (4)$$

$$\begin{cases} \Delta B(j) = RUE (1 - e^{-K \cdot LAI(j)}) I(j) & TT(j) \leq TT_M \\ \Delta B(j) = 0 & TT(j) > TT_M \end{cases} \quad (5)$$

$$\begin{cases} \Delta LAI(j) = \alpha \Delta TT(j) LAI(j) \max[LAI_{max} - LAI(j), 0] & TT(j) \leq TT_L \\ \Delta LAI(j) = 0 & TT(j) > TT_L \end{cases} \quad (6)$$

Les variables d'entrée, à savoir les températures minimale et maximale TMIN(j), TMAX(j) et l'intensité du rayonnement solaire I(j) apparaissent explicitement dans les équations (3) à (5). Les paramètres sont T_{base} (la température de base assurant la croissance), RUE (l'efficacité d'utilisation du rayonnement solaire), K (coefficient d'extinction de la lumière), α (le taux d'accroissement de l'index foliaire pour de faibles valeurs de LAI) et le LAI_{max} (l'index foliaire maximum). Dans ces équations, TT_M et TT_L sont deux seuils au-delà desquels, respectivement, la biomasse et le LAI cessent de croître.

À côté de la représentation d'état, le modèle de culture peut se présenter sous la forme d'un modèle « entrée – sortie ». Celui-ci s'obtient en intégrant pas à pas les équations aux différences (2) dans lesquelles les variables d'état initiales ont été introduites ainsi que les variables d'entrée actualisées. Dans la majorité des cas, on ne s'intéresse pas à tous les résultats que le modèle est capable de fournir, mais seulement à

la valeur finale de certains d'entre eux (qui sont des variables d'état ou des fonctions de celles-ci), comme par exemple le prélèvement d'azote par les plantes, l'azote lessivé vers le sol, la biomasse à un moment donné, le stress hydrique à un moment donné, etc. Ces variables de réponse sont notées $Y(t)$ et le modèle prend alors la forme suivante :

$$Y(t) = f(U(t), X(t), \theta) \quad (7)$$

On peut donc représenter le modèle de culture selon le schéma-bloc donné à la **figure 1**, dans lequel les variables d'état issues des équations aux différences apparaissent comme des variables internes au système.

Le choix des variables d'entrée et des paramètres du modèle implique de prendre des décisions sur les éléments importants de la dynamique du système. Le fait d'ajouter une nouvelle variable dans les équations apporte théoriquement plus de précision dans la réponse du système, mais requiert une estimation paramétrique supplémentaire, ce qui risque de conduire à des erreurs additionnelles et peut, paradoxalement, conduire à une diminution de la précision des prédictions.

Par ailleurs, un certain nombre d'incertitudes existent dans les modèles de culture. Dans les modèles déterministes, elles ne sont pas prises en compte. On peut donc considérer que le modèle déterministe décrit le comportement moyen du système. Pour prendre en compte les incertitudes du modèle et des mesures, la représentation d'état (2) est exprimée sous forme d'un système :

$$\begin{cases} U(t+\Delta t) = U(t) + g[U(t), X(t), \theta] + \eta(t) \\ M(t) = C \cdot U(t) + \tau(t) \end{cases} \quad (8)$$

avec $M(t)$ le vecteur des mesures et C la matrice des valeurs mesurées. $\eta(t)$ et $\tau(t)$ sont des variables aléatoires, généralement considérées comme ayant une distribution normale de moyenne nulle. $\eta(t)$ représente les erreurs de modélisation et les erreurs sur les variables d'entrée, tandis que $\tau(t)$ correspond aux erreurs de mesure. Les erreurs de modélisation $\eta(t)$ ont trois origines :

- les équations décrivent l'évolution de phénomènes biologiques et physiques par des approximations (par exemple, relations empiriques établies dans un

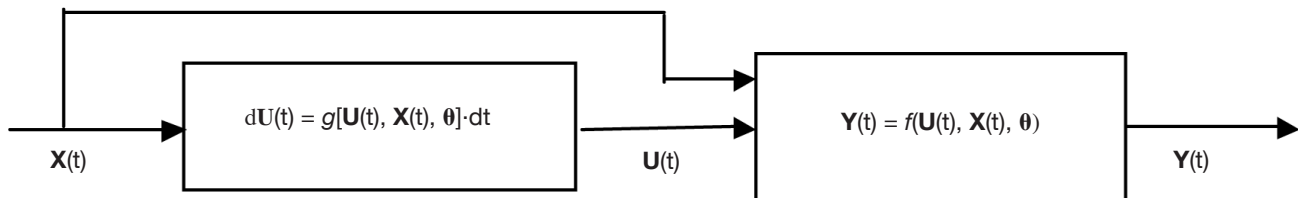


Figure 1. Représentation schématique d'un modèle de culture — *Schematic structure of a crop simulation model.*

contexte plus restrictif que celui dans lequel le modèle est utilisé),

- des erreurs existent sur l'estimation des paramètres,
- les variables d'entrée peuvent être entachées d'erreur.

Weiss et al. (2006) citent à titre d'exemple la dérive de capteurs équipant une station météorologique dont la maintenance n'est pas assurée avec une fréquence suffisante. Les erreurs de mesure $\tau(t)$ englobent les erreurs liées aux observations des variables d'état. Ainsi, on peut trouver des erreurs sur la mesure du rendement en grain effectuée de manière manuelle ou à l'aide de machines de récolte, des erreurs d'appréciation du stade de développement d'une céréale, etc. Dans certains cas, on néglige les erreurs de mesure et les incertitudes sont uniquement prises en compte par l'équation :

$$\mathbf{U}(t + \Delta t) = \mathbf{U}(t) + g[\mathbf{U}(t), \mathbf{X}(t), \boldsymbol{\theta}] + \boldsymbol{\eta}(t) \quad (9)$$

Si le modèle est représenté sous forme des relations entrée – sortie, la prise en compte des incertitudes se fait en introduisant un terme d'erreur dans l'équation (7) :

$$\mathbf{Y}(t) = f(\mathbf{U}(t), \mathbf{X}(t), \boldsymbol{\theta}) + \boldsymbol{\epsilon}(t) \quad (10)$$

La difficulté majeure dans l'évaluation de $\boldsymbol{\epsilon}(t)$ est liée au fait que les erreurs associées aux variables de sortie peuvent présenter des variances différentes et être corrélées entre elles (Makowski et al., 2001b). Selon

Makowski et al. (2006), c'est le cas de l'estimation des valeurs de biomasse réalisées au cours d'une saison culturale pour chaque couple site – année d'une expérimentation.

3. LES CRITÈRES DE VALIDATION DES MODÈLES

La validation d'un modèle est une étape cruciale, comme en attestent les nombreuses définitions existant au travers de la littérature. Le **tableau 1**, adapté de Bellochi et al. (2009), reprend quelques-unes de ces définitions. D'une manière générale, un modèle de culture n'est une bonne représentation de la réalité que s'il peut être utilisé pour prédire un phénomène observable, à l'intérieur de la gamme pour laquelle il est calibré, avec suffisamment d'exactitude (Loague et al., 1991). La mise en œuvre de cette assertion implique de définir les critères qui permettent de considérer qu'un modèle est « acceptable » en fonction des objectifs qu'il poursuit et ensuite de le tester en fonction de ces critères. En première approche vient la comparaison, au moyen d'un test statistique, des moyennes des valeurs mesurées et simulées, accompagnées de leur écart-type. Les performances d'un modèle pourront être jugées acceptables s'il n'est pas possible de rejeter l'hypothèse de non-différence des valeurs observées et prédites. En utilisant ce type de test, deux types d'erreurs majeures peuvent se produire :

Tableau 1. Définitions de la validation d'un modèle (adapté de Bellochi et al., 2009) — *Definition of the validation of a model (adapted from Bellochi et al., 2009).*

| Définition | Source |
|---|---|
| “A model is a good representation of the reality only if it can be used to predict, within a calibrated and validated range, an observable phenomenon, with acceptable accuracy and precision” | Loague et al., 1991 |
| “The validity of a crop model should be judged by the soundness of the hypothesis on which it is based and on logical structure. It should also meet a “behavioral test” that indicates whether the model subcomponents behave, at least qualitatively, in accordance with observation” | Sinclair et al., 2000 |
| “A procedure consisting in comparing the model output with field or laboratory data to prove the model efficiency” | Dias et al., 2006 ¹ cité par Bellochi et al., 2009 |
| “A procedure consisting in verifying if the model is able to reproduce data, independently of those involved in its calibration” | Cardoso et al., 2008 ² cité par Bellochi et al., 2009 |
| “The assessment of the performance of a model against an independently collected dataset” | Matthews et al., 2008 ³ cité par Bellochi et al., 2009 |

¹Dias J.M. & Lopes J.F., 2006. Implementation and assessment of hydrodynamic, salt and heat transport models: the case of Ria de Aveiro Lagoon (Portugal). *Environ. Modell. Softw.*, **21**, 1-15 ; ² Cardoso A.C. & Lopes J.F., 2008. 3D ecological modelling of the Aveiro coast (Portugal). In: Sánchez-Marrè M. et al., eds. *Integrating sciences and information technology for environmental assessment and decision making. Proceedings of the 4th Biennial Meeting of the International Environmental Modelling and Software Society, 7-10 July, Barcelona, Spain*. Vol. 1, 181-190 ; ³ Matthews K.B. et al., 2008. Raising the bar - Is evaluating the outcomes of decision and information support tools a bridge too far? In: Sánchez-Marrè M. et al., eds. *Integrating sciences and information technology for environmental assessment and decision making. Proceedings of the 4th Biennial Meeting of the International Environmental Modelling and Software Society, 7-10 July, Barcelona, Spain*. Vol. 1, 948-955.

- erreur de type I : c'est l'erreur portant sur le développement du modèle, elle correspond au rejet d'une hypothèse vraie. Les équations constituant le modèle sont exactes, mais le modèle est mal paramétré, donc il lui est impossible de simuler correctement le phénomène ;
- erreur de type II : c'est l'erreur portant sur l'utilisation du modèle, elle correspond à l'acceptation d'une hypothèse fautive : le modèle est faux / incomplet, mais permet toutefois de simuler de manière acceptable le phénomène dans la gamme de valeurs pour laquelle il a été paramétré.

C'est pourquoi un second type d'évaluation des performances est souvent considéré, à savoir l'analyse des erreurs résiduelles entre les valeurs simulées et observées. Une synthèse des principaux critères utilisés dans l'analyse des discordances entre mesures et simulations est faite ici. Il s'agit des critères les plus couramment employés en simulation agronomique. Une liste plus exhaustive de critères a été établie par Loague et al. (1991) ou encore par Wallach et al. (2006). Le critère le plus simple pour mesurer les performances d'un modèle consiste à calculer, pour chaque couple i de valeurs mesurées et modélisées, la différence entre la variable mesurée Y_i et la valeur correspondant \hat{Y}_i estimée par le modèle :

$$D_i = Y_i - \hat{Y}_i \quad (11)$$

Les valeurs de D_i peuvent être synthétisées en calculant leur moyenne, connue sous le nom de biais du modèle :

$$\text{Biais} = \frac{1}{N} \sum_{i=1}^N D_i \quad (12)$$

où N est le nombre total de couples de valeurs correspondant à un même cas d'étude, d'une part mesuré, d'autre part simulé. Le biais n'est pas une mesure suffisante de la qualité d'un modèle. En effet, un faible biais peut traduire de très faibles erreurs dans toutes les situations ou peut au contraire résulter de grandes erreurs qui se compensent mutuellement. La racine carrée de l'erreur quadratique moyenne (RMSE, *Root mean square error*) permet d'éliminer ce dernier problème :

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (D_i)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \quad (13)$$

L'avantage de la RMSE est qu'elle s'exprime dans les mêmes unités que Y_i , ce qui facilite son interprétation. Cependant, la mise au carré de D_i donne

un poids plus important aux erreurs les plus grandes. Il convient donc d'être vigilant lors de l'interprétation car une grande RMSE peut être le fruit d'une ou deux grandes différences seulement. La RMSE peut être exprimée sous forme relative en la divisant par la moyenne des valeurs observées, le quotient est appelé coefficient moyen de variation :

$$\text{RRMSE} = \frac{\text{RMSE}}{\bar{Y}} \quad (14)$$

L'erreur maximale est donnée par :

$$\text{ME} = \text{Max} |Y_i - \hat{Y}_i|_{i=1}^n \quad (15)$$

L'efficacité de modèle et le coefficient de corrélation présentent une borne inférieure et/ou supérieure, ce qui facilite leur interprétation et les rend pertinents pour comparer des modèles. L'efficacité du modèle est donnée par :

$$\text{EF} = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \quad \text{EF} \leq 1 \quad (16)$$

Si le modèle est parfait, alors $Y_i = \hat{Y}_i$ pour toutes les valeurs de i et $\text{EF} = 1$. De par sa forme mathématique, l'efficacité de modèle permet de juger la performance globale du modèle par rapport à une prédiction qui serait simplement égale à la moyenne des observations. Le coefficient de corrélation de Pearson est donné par :

$$r = \frac{\sigma_{Y\hat{Y}}}{\sigma_Y \sigma_{\hat{Y}}} \quad -1 \leq r \leq 1 \quad (17)$$

où σ_Y^2 , $\sigma_{\hat{Y}}^2$, $\sigma_{Y\hat{Y}}$ sont respectivement les variances de Y , de \hat{Y} et la covariance de Y et \hat{Y} . Une valeur $r = 1$ indique une relation parfaitement linéaire entre Y_i et \hat{Y}_i , mais ne signifie pas que les valeurs prédites sont proches des valeurs observées. Supposons par exemple que $\hat{Y}_i = 0,5 Y_i$ pour toutes les valeurs de i . Le coefficient de corrélation $r = 1$, alors que le modèle prédit systématiquement des valeurs qui sont la moitié des valeurs observées.

La déviation normalisée (ND), parfois trouvée dans la littérature sous l'appellation de biais moyen (MB, *Mean bias*), peut être positive ou négative, mais est idéalement nulle :

$$ND = \frac{\sum_{i=1}^N Y_i - \sum_{i=1}^N \hat{Y}_i}{\sum_{i=1}^N Y_i} \quad (18)$$

Les critères donnés aux équations (11) à (18) sont rarement utilisés seuls pour évaluer la qualité d'un modèle. Brisson et al. (2002) et Beaudoin et al. (2008) utilisent conjointement RMSE, EF, ND et considèrent que la calibration ou la validation est adéquate si :

$$\begin{aligned} EF &\geq 0,5 \\ |ND| &\leq 0,1 \end{aligned} \quad (19)$$

Les critères les plus couramment utilisés, leurs gammes de valeurs, les valeurs idéales et les valeurs utilisées en pratique sont synthétisées au **tableau 2**. Par ailleurs, il est possible de décomposer le RMSE entre d'une part l'erreur quadratique moyenne systématique $RMSE_s$ et d'autre part, l'erreur quadratique moyenne non systématique $RMSE_u$, comme l'ont fait Houllès et al. (2004). Cette décomposition permet de faire la distinction entre la composante systématique de l'erreur due à un biais du modèle et une erreur aléatoire :

$$RMSE_s = \sqrt{\frac{1}{N} \sum_{i=1}^N (b + aY_i - \hat{Y}_i)^2} \quad (20)$$

$$RMSE_u = \sqrt{\frac{1}{N} \sum_{i=1}^N (b + aY_i - \hat{Y}_i)^2} \quad (21)$$

avec a et b respectivement la pente et l'ordonnée à l'origine de la droite de régression de \hat{Y}_i en fonction de Y_i .

Tableau 2. Critères utilisés pour juger de la qualité des modèles — *Criteria to judge the quality of the models.*

| Critère | Gamme de valeurs | Valeurs idéales | Valeurs pratiques |
|---------|-----------------------|-----------------|-------------------|
| Di | $[-\infty ; +\infty]$ | 0 | |
| Biais | $[-\infty ; +\infty]$ | 0 | |
| RMSE | $[0 ; +\infty]$ | 0 | |
| RRMSE | $[0 ; +\infty]$ | 0 | |
| ME | $[0 ; +\infty]$ | 0 | |
| EF | $[-\infty ; 1]$ | 1 | $EF \geq 0,5$ |
| r | $[-1 ; +1]$ | ± 1 | |
| ND | $[-\infty ; +\infty]$ | 0 | $ ND \leq 0,1$ |

Notons que, dans les modèles écophysiologiques, il est souvent possible de réduire l'erreur due à un biais du modèle par correction des formalismes mathématiques du modèle ou par optimisation de la valeur des paramètres, alors que l'erreur non systématique, par exemple liée à la variabilité génétique naturelle d'un peuplement, ne peut être réduite par l'amélioration du modèle. Il convient toutefois d'insister sur le fait qu'un nombre plus important de mesures (collecte d'un plus grand nombre d'échantillons à une fréquence accrue et pour chaque site – année) permet de réduire la variabilité de l'échantillonnage. La stabilisation de la moyenne des mesures permet alors d'assurer l'obtention d'une $RMSE_u$ plus exacte.

4. L'ESTIMATION PARAMÉTRIQUE

L'objectif de l'estimation paramétrique est d'évaluer les paramètres θ du modèle en partant d'un échantillon de variables observées (Flenet et al., 2004 ; Launay et al., 2005). La complexité des modèles de culture est telle qu'en général, plus d'une centaine de paramètres est nécessaire. D'un autre côté, les variables d'entrée et d'état mesurées sont souvent limitées par la nécessité de réaliser de longues expérimentations pendant plusieurs saisons culturales, avec des variétés différentes, ce qui représente un coût important en termes financiers et en ressources humaines. Il ne faut toutefois pas perdre de vue que la qualité et la quantité de ces mesures influencent directement la qualité de l'estimation paramétrique. Cette dernière reste toujours un compromis entre l'obtention d'un modèle robuste et générique ou d'un modèle spécifique et local.

4.1. Incidence des mesures sur l'estimation paramétrique

Étant donné que les performances prédictives d'un modèle sont liées à la qualité de sa paramétrisation, de nombreux auteurs insistent sur la difficulté de paramétrer correctement un modèle de culture et mettent l'accent sur les différentes sources d'erreurs qui y sont liées.

Makowski et al. (2001a et 2002) pensent que l'estimation des paramètres des modèles fait en général l'objet de trop peu d'attention. Par exemple, dans des essais portant sur les niveaux de fumure, plusieurs mesures sont généralement réalisées sur le même site, au cours d'une même année, avec différentes modalités de fertilisation. Les auteurs estiment que, dans ce cas, il n'est pas raisonnable de supposer que les erreurs sur les mesures sont indépendantes pour les mêmes combinaisons site-année et qu'il convient d'en tenir compte. Dans une étude identifiant les sources de réduction du rendement, Lawless et al. (2008) montrent

l'importance de prendre en compte la variabilité spatiale des propriétés hydrauliques des sols. En effet, la teneur en eau du sol affecte non seulement le bilan hydrique, mais également la disponibilité en azote, qui elle-même influe à son tour le rendement, tout particulièrement au cours des périodes de sécheresse. À partir d'une étude de sensibilité, Ruget et al. (2002) montrent que les paramètres décrivant le système aérien ont un effet sur le rendement et le prélèvement d'eau, mais influencent peu la lixiviation d'azote. Par contre, les paramètres décrivant le sol (en particulier, la capacité au champ) et la zone racinaire ont un effet important sur l'eau drainée et la lixiviation d'azote. Beaudoin et al. (2008) insistent sur l'importance de réaliser l'estimation paramétrique à partir de mesures effectuées sur de longues périodes. Ils affirment que la capacité des modèles à prédire à long terme les pertes en eau et en nitrates doit rester au cœur des préoccupations, notamment du fait que les modèles de culture ont, jusqu'à présent, été majoritairement évalués sur des expériences annuelles. Basso et al. (2010), qui étudient la réponse du blé à la fertilisation azotée, signalent que quelques années de mesures expérimentales pourraient ne pas refléter la réponse d'une culture, principalement en raison de la grande variabilité des conditions climatiques et, en particulier, des précipitations durant la période de culture. Finalement, selon Jongschaap (2007), un grand nombre de mesures est nécessaire pour réaliser l'estimation paramétrique. En outre, celle-ci doit être effectuée dans les conditions où les données ont été collectées pour générer des résultats satisfaisants. Même dans ce cas, les résultats de simulations peuvent dévier des observations, en particulier si les données d'entrée sont difficiles à mesurer avec précision. Dans de telles situations, on peut se voir contraint de remplacer les paramètres qui auraient dû être optimisés par des informations fournies par des experts ou par des données issues de la littérature.

Au vu de la complexité et de l'interaction des phénomènes intervenant dans les modèles de culture, il convient donc de définir des stratégies de modélisation adéquates, présentant un bon compromis entre d'une part, la quantité d'informations nécessaire pour réaliser l'estimation paramétrique (nombre de variables étudiées au cours d'une même année, nombre de répétitions, nombre de mesures en cours de saison, nombre d'années de mesure, etc.) et d'autre part, les coûts représentés par l'expérimentation de terrain.

4.2. Méthodes d'estimation paramétrique

Les méthodes d'estimation paramétrique sont basées sur le fait que les variables observées sont des réalisations de variables aléatoires, dans la mesure où un essai agronomique n'est jamais parfaitement reproductible. L'estimation de paramètres n'est pas un problème

simple à résoudre dans les modèles de culture : un grand nombre d'équations sont non linéaires, le nombre de paramètres est élevé et les variabilités spatiales (hétérogénéité des parcelles) et temporelles (fluctuations climatiques) sont grandes. Les théories statistiques fournissent toutefois différentes méthodes permettant d'extraire des données les informations reproductibles. On distingue l'approche fréquentielle et l'approche bayésienne.

L'approche fréquentielle se base sur un échantillon de données pour trouver la valeur vraie des paramètres en considérant que ceux-ci ne sont pas aléatoires. Elle repose donc sur le fait qu'un set unique de paramètres satisfait à l'optimisation du modèle. Dans l'espace paramétrique multi-dimensionnel, l'approche fréquentielle reste cependant sensible aux minima locaux du critère d'optimisation (par exemple, minimisation de la RMSE). On peut distinguer trois variantes. La première consiste à trouver le maximum de paramètres à partir de différentes sources d'information (mesures, littérature, fonctions de pédotransfert, etc.) et à estimer les paramètres restant par un ajustement statistique. La deuxième variante consiste à utiliser une méthode de régression séquentielle (« *stepwise regression* ») pour choisir successivement les paramètres qui conduisent au meilleur ajustement du modèle (Wallach et al., 2001 ; Varella et al., 2010 ; Wallach et al., 2011). Enfin, la troisième variante consiste à mener une analyse de sensibilité pour estimer les paramètres qui ont le plus d'influence sur les réponses du modèle (Ruget et al., 2002). Dans tous les cas se pose le problème du nombre de paramètres à estimer. En particulier, il faut éviter le surajustement (« *overfitting* ») qui consiste à ajuster un trop grand nombre de paramètres par rapport aux données disponibles et aux incertitudes qui les affectent. Différents cas de surajustement peuvent se présenter. Le premier et le plus classique consiste à chercher à optimiser deux paramètres intervenant dans une même équation. Dans ce cas, il convient de fixer l'un d'entre eux à une valeur nominale. Un autre cas de surajustement se rencontre lorsque trop peu de données sont disponibles. Tout comme il est impossible de résoudre un système d'équations où le nombre d'inconnues est supérieur au nombre d'équations, il est impossible d'optimiser plus de paramètres qu'il y a de données disponibles. L'emploi des méthodes fréquentielles ne permet donc en général pas d'optimiser tous les paramètres simultanément (Bechini et al., 2006).

Les méthodes bayésiennes sont de plus en plus utilisées pour estimer les paramètres de modèles complexes. Elles les estiment à partir de deux types d'information : un échantillon de données (comme dans la méthode fréquentielle) et une information *a priori* concernant les paramètres. Dans un premier

temps, on définit donc une distribution de probabilité *a priori* du paramètre recherché, basée sur la littérature ou l'expertise. Dans un deuxième temps, on calcule une distribution de probabilité *a posteriori* à partir de la distribution précédente et des données disponibles en utilisant le théorème de Bayes et la définition d'une fonction de vraisemblance, cette dernière permettant de comparer les simulations aux mesures. Le paramètre est alors déterminé comme l'espérance mathématique ou le mode principal de la distribution *a posteriori*. Dans la mesure où la méthode bayésienne utilise deux types d'information, elle est parfois classée dans les méthodes d'assimilation de données (voir ci-dessous).

Traditionnellement, l'estimation de la distribution *a posteriori* s'avérait compliquée, sauf dans des cas simples. Avec l'évolution de la vitesse de calcul des CPU, cette estimation connaît un regain d'intérêt. Une première manière de procéder réside dans l'estimation de l'incertitude sur la fonction de vraisemblance (GLUE, *Generalised Likelihood Uncertainty Estimation*). Cette méthode permet de traiter les problèmes pour lesquels les hypothèses simples sur les fonctions de vraisemblance ne sont pas d'application (Vrugt et al., 2009). Elle réalise l'optimisation des paramètres par une discrétisation de l'espace paramétrique. Elle peut cependant conduire à une représentation inexacte de la distribution *a posteriori* des paramètres lorsque leur nombre est important (Makowski et al., 2002).

Une approche plus sophistiquée pour évaluer à la fois la distribution *a posteriori* des paramètres et l'incertitude sur la prédiction des modèles réside dans l'emploi des chaînes de Markov-Monte Carlo (MCMC, *Markov Chain Monte Carlo*) (Metropolis et al., 1953 ; Vrugt et al., 2009). En effet, les simulations MCMC se basent sur une représentation formelle de la fonction de vraisemblance, ce qui permet d'échantillonner correctement les régions de l'espace paramétrique multidimensionnel et de séparer les régions de haute densité de probabilité des autres solutions dites « non comportementales ». En 1970, Hastings a étendu la théorie MCMC pour inclure tout type de distributions de probabilité. Cette extension, appelée l'algorithme de *Metropolis Hastings* (MH), est à l'origine de la plupart des approches bayésiennes formelles actuelles.

5. L'ASSIMILATION DE DONNÉES

L'assimilation de données a pour objectif d'intégrer dans la modélisation des sources d'information diverses (observations, mesures, etc.) sur les variables d'état du système afin de construire un état du système « plus vrai » que l'état « simulé » obtenu par la modélisation classique. Ces méthodes, largement utilisées en météorologie, n'en sont qu'à leurs débuts

dans les modèles de culture. Elles sont l'analogie des principes de filtrage appliqués dans les domaines de l'automatique et de la dynamique des systèmes. Elles reposent sur l'hypothèse selon laquelle la trajectoire suivie par un objet (l'évolution temporelle de la biomasse, par exemple) est connue de manière théorique à chaque instant (ici, chaque jour) et est à même d'atteindre une cible (la mesure du phénomène observé). Cependant, suite à l'existence de diverses perturbations (erreurs de mesure des variables d'entrée, incident climatique, maladie, etc.), il est possible que le modèle dévie de la « bonne trajectoire ». Il convient alors d'établir des balises (mesures en cours de saison) qui permettront de corriger la simulation lorsque le modèle dévie exagérément de sa cible.

D'une manière générale, on distingue deux types d'approches dans les méthodes d'assimilation de données. La première est basée sur la comparaison des valeurs d'état simulées et mesurées, chaque fois que celles-ci sont disponibles. Si l'écart entre les deux grandeurs est trop important, la variable d'état simulée est remplacée par une valeur proche de la valeur mesurée, calculée par filtrage. Les simulations sont alors relancées depuis ce point précis, avec la variable d'état actualisée.

La deuxième approche conduit à opérer une correction des paramètres (qui ne sont donc pas fixés une fois pour toutes) en cours de simulation. Dans cette démarche, on distingue deux modes d'action. Le premier consiste à réévaluer les performances du modèle chaque fois qu'une nouvelle mesure est disponible et à voir si une autre valeur du paramètre étudié permet d'améliorer les performances prédictives sur la gamme simulée. La valeur retenue du paramètre est alors utilisée pour réaliser les nouvelles simulations. Le second mode d'action implique de considérer que les paramètres peuvent évoluer au cours du temps (d'une année à l'autre, ou entre deux stades végétatifs, par exemple) et à ajuster le paramètre étudié pour simuler au mieux l'évolution de la variable d'état uniquement dans la plage temporelle comprise entre deux mesures consécutives. Le paramètre étudié peut ainsi être amené à évoluer pour chaque couple de mesures consécutives. C'est sa stabilité temporelle, ou non, qui justifiera le fait de pouvoir le considérer comme un paramètre fixe ou davantage comme une variable interne du système.

Le choix de l'approche dépend de la question à laquelle l'utilisateur est amené à répondre. Faut-il « mettre à jour » les variables d'état ou les paramètres du modèle ? Choisir de n'adapter que les variables d'état revient à considérer que le comportement du champ est celui d'un champ moyen. Les paramètres du modèle correspondent à ceux d'un champ moyen, mais pour une raison spécifique, le modèle n'est pas à même de simuler la réalité. Il est alors décidé de filtrer uniquement les variables de sortie du modèle. En

général, le modèle ne doit être filtré que sur quelques couples de valeurs mesures – simulations d'une variable d'état. La raison de ce choix s'explique, par exemple, par l'existence d'un incident climatique que le modèle ne peut prendre en compte ou par une erreur de mesure sur les variables de forçage du modèle.

Dans le cas contraire, si l'on décide d'optimiser les paramètres du modèle pour chaque nouvelle mesure disponible, l'hypothèse retenue par l'utilisateur est que le comportement de la culture est différent de celui d'un champ moyen. Un tel exemple de situation pourrait être rencontré en sélection génétique, lorsqu'on cherche à sélectionner de nouveaux cultivars, présentant des caractéristiques particulières (précocité, sensibilité réduite au stress, etc.). Certaines caractéristiques du nouveau cultivar sont différentes de celles de la culture initiale, alors que d'autres caractères sont conservés. Lors de la réalisation de simulations sur cette nouvelle culture, les paramètres doivent être modifiés pour rendre compte des particularités dans la gamme temporelle pour laquelle elles s'expriment et durant laquelle des mesures ont été réalisées. Dans ce cas, il ne faut plus filtrer les sorties du modèle, mais bien mettre à jour les paramètres.

Jongschaap (2007) a ainsi étudié le rendement de maïs à l'aide du modèle PlantSys en corrigeant les résultats de la simulation, par l'introduction dans le modèle en temps réel du LAI et du contenu en azote de la canopée (*LeafNwt*). Il estime que ces variables sont importantes dans la mesure où elles conditionnent l'interception de la lumière et la capacité de production photosynthétique. Elles peuvent être estimées à partir de données issues de la télédétection, acquises de préférence au moment de la phase de reproduction. Naud et al. (2008) cherchent à estimer le statut azoté du blé d'hiver à l'aide du modèle Azodyn. Ils utilisent une approche bayésienne appelée « filtres particuliers » pour corriger trois variables d'état, à savoir la biomasse aérienne, le prélèvement d'azote et la quantité d'azote accumulée dans le sol. Ils concluent que les résultats obtenus avec une seule correction, de préférence réalisée à partir du stade 31 sur l'échelle de Zadoks (Zadoks et al., 1974) (premier nœud détectable), sont satisfaisants. De Wit et al. (2007) utilisent un « filtre de Kalman d'ensemble » pour intégrer des mesures de teneurs en eau obtenues par voie satellitaire dans le modèle de culture WOFOST. Les résultats des simulations ne sont pas parfaitement satisfaisants, ce qui laisse supposer que le filtre n'est pas optimisé. Yuping et al. (2008) couplent le modèle SAIL avec le modèle WOFOST, *via* la variable d'état LAI. Ils constatent que le modèle WOFOST donne des estimations plus précises et plus justes de la biomasse aérienne lorsque des données satellitaires permettent de réinitialiser les simulations de LAI en sortie d'hiver. La procédure appliquée est donc ici celle du filtrage d'une variable de sortie, sur la première mesure de

sortie d'hiver, mais en actualisant la valeur du LAI, obtenue *via* l'imagerie satellitaire et le modèle SAIL. Cette amélioration des simulations est principalement expliquée par une mauvaise prise en compte au sein du modèle des pertes de culture en hiver.

6. CONCLUSION

La modélisation de la croissance et du développement des plantes cultivées est complexe, aussi bien au niveau des formalismes mathématiques utilisés qu'au niveau de la paramétrisation.

Pour s'assurer que la paramétrisation est correcte, l'utilisateur doit s'imposer des critères permettant de juger d'une part que le modèle est calibré correctement et, d'autre part, qu'il possède une réelle capacité prédictive. Le critère le plus employé est la RMSE, à la fois pour sa simplicité et sa fiabilité. D'autres critères souvent utilisés conjointement sont la déviation normalisée et l'efficacité du modèle qui permettent respectivement de juger directement de la sur- ou sous-estimation systématique du modèle et de juger de la supériorité du modèle par rapport à l'emploi d'une simple moyenne des observations passées du phénomène.

Dès que ces critères sont imposés, il y a lieu de procéder à un choix des paramètres à optimiser. De manière générale, il convient d'adapter le(s) paramètre(s) auquel(s) on s'intéresse aux mesures disponibles. Inversement, il faut adapter les mesures à réaliser au(x) paramètre(s) que l'on cherche à optimiser. Ainsi, des paramètres contrôlant la croissance d'une culture seront optimisés sur des années ne présentant aucun stress et des paramètres de sensibilité seront optimisés dans la situation contraire.

En ce qui concerne les méthodes d'optimisation, on distingue les méthodes fréquentielles et les méthodes bayésiennes. Les premières présentent un avantage certain en temps de calcul lorsque seuls quelques paramètres doivent être optimisés et/ou lorsque beaucoup de données sont disponibles. Par ailleurs, elles sont également intéressantes lorsque le paramètre a une signification claire pour l'utilisateur (par exemple, un paramètre décrivant la somme de température entre deux stades physiologiques). L'expertise de l'utilisateur est importante car elle lui permet de juger si la valeur d'un paramètre est plausible ou, au contraire, aberrante. Finalement, ces méthodes sont également intéressantes dans la mesure où l'utilisateur accorde une grande confiance aux variables d'entrée du modèle.

Les méthodes bayésiennes présentent un avantage considérable lorsque beaucoup de paramètres doivent être optimisés et/ou que peu de mesures sont disponibles. Elles permettent également à l'utilisateur de juger de l'incertitude induite par les paramètres sur la

réponse finale du modèle. Pour les problèmes relevant d'une étude de sensibilité des sorties d'un modèle, ces méthodes devraient être privilégiées. En effet, une propriété implicite de l'approche bayésienne permet d'affirmer qu'au plus un paramètre estimé présente un mode marqué, au plus la sortie du modèle est sensible à ce paramètre. Il est également intéressant de coupler l'incertitude liée à une variable d'entrée à celle d'un paramètre du modèle et d'étudier leur interaction sur la réponse finale du modèle.

Les méthodes d'assimilation de données constituent une voie intéressante, mais leur efficacité requiert une bonne paramétrisation préalable du modèle. Leur intérêt réside surtout dans la prise en compte de phénomènes non prévisibles ou se situant momentanément en-dehors de la gamme de paramétrisation (incidents climatiques, etc.), dans les cas où certaines variables d'entrée ne sont pas connues avec suffisamment de précision ou pour lesquelles l'incertitude est élevée. Leur application est également un atout lors d'une modification soudaine de la variable à simuler, pour une raison indépendante et non maîtrisée par l'utilisateur (comme, par exemple, la prise en compte de maladies). Par ailleurs, en poussant les techniques de filtrage plus loin que la seule amélioration des sorties du modèle, c'est-à-dire en optimisant à la fois ces sorties et certains paramètres simultanément, il est possible d'arriver à considérablement améliorer les résultats des simulations.

Remerciements

Les auteurs de l'article adressent leurs remerciements au Service public de Wallonie (DGARNE) pour le financement du projet intitulé « Suivi en temps réel de l'environnement d'une parcelle agricole par un réseau de microcapteurs en vue d'optimiser l'apport en engrais azotés ».

Bibliographie

- Basso B. & Ritchie J.T., 2005. Impact of animal manure, compost and inorganic fertilizer on nitrate leaching and yield in a six-year maize alfalfa rotation. *Agric. Ecosyst. Environ.*, **108**, 329-341.
- Basso B. et al., 2009. Landscape position and precipitation effects on spatial variability of wheat yield and grain protein in Southern Italy. *J. Agron. Crop Sci.*, **195**, 301-312.
- Basso B. et al., 2010. Long-term wheat response to nitrogen in a rainfed Mediterranean environment: field data and simulation analysis. *Eur. J. Agron.*, **33**, 132-138.
- Beaudoin N. et al., 2008. Evaluation of the soil crop model STICS over 8 years against the "on farm" database of Bruyères catchment. *Eur. J. Agron.*, **29**, 46-57.
- Bechini L., Bocchi S., Maggiore T. & Confalonieri R., 2006. Parameterization of a crop growth and development simulation model at sub model component level. An example for winter wheat (*Triticum aestivum* L.). *Environ. Modell. Softw.*, **21**, 1042-1054.
- Bellochi G., Rivington M., Donatelli M. & Matthews K., 2009. Validation of biophysical models: issues and methodologies. A review. *Agron. Sustainable Dev.*, **30**, 109-130.
- Brisson N. et al., 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. Theory and parameterization applied to wheat and corn. *Agronomie*, **18**, 311-346.
- Brisson N. et al., 2002. STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. *Agronomie*, **22**, 69-82.
- Brisson N. et al., 2003. An overview of the crop model STICS. *Eur. J. Agron.*, **18**, 309-332.
- Brisson N., Launay M., Mary B. & Beaudoin N., eds, 2008. *Conceptual basis, formalisations and parameterization of the STICS crop model*. Versailles, France: Éditions Quæ.
- De Wit A.J.W. & Van Diepen C.A., 2007. Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts. *Agric. For. Meteorol.*, **146**, 38-56.
- Flenet F., Villon P. & Ruget F., 2004. Methodology of adaptation of the STICS model to a new crop: spring linseed (*Linum usitatissimum* L.). *Agronomie*, **24**, 367-381.
- Génard M., Jeuffroy M.H., Jullien A. & Quilot B., 2006. *Les modèles en écophysologie pour l'action en agronomie et la création variétale*. Versailles, France: Éditions Quæ.
- Hansen S., Jensen H.E., Nielsen N.E. & Swenden H., 1990. *Daisy, a soil plant system model. Danish simulation model for transformation and transport of energy and matter in the soil plant atmosphere system*. Copenhagen: The National Agency for Environment Protection.
- Hastings W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97-109.
- Houlès V. et al., 2004. Evaluation of the ability of the crop model STICS to recommend nitrogen fertilisation rates to agro-environmental criteria. *Agronomie*, **24**, 339-349.
- Jones C.A. & Kiniry J.R., 1986. *CERES-Maize, a simulation model of maize growth and development*. College Station, TX, USA: Texas A&M University Press.
- Jongschaap R.E.E., 2007. Sensitivity of a crop growth simulation model to variation in LAI and canopy nitrogen used for run-time calibration. *Ecol. Modell.*, **200**, 89-98.
- Launay M., Flenet F., Ruget F. & Garcia de Cortazar Atauri I., 2005. Généricité et méthodologie d'adaptation de STICS à de nouvelles cultures. In : *Actes du Séminaire STICS, mars 2005, Carry-le-Rouet, France*.
- Lawless C., Semenov M.A. & Jamieson P.D., 2008. Quantifying the effect of uncertainty in soil moisture

- characteristics on plant growth using a crop simulation model. *Field Crops Res.*, **106**(2), 138-147.
- Loague K. & Green R.E., 1991. Statistical and graphical method for evaluating solute transport models: overview and application. *J. Contam. Hydrol.*, **7**, 51-73.
- Loyce C. & Wery J., 2006. *Les outils des agronomes pour l'évaluation et la conception des systèmes de culture*. Versailles, France : Éditions Quæ.
- Makowski D. & Wallach D., 2001a. How to improve model-based decision rules for nitrogen fertilization. *Eur. J. Agron.*, **15**(3), 197-208.
- Makowski D., Wallach D. & Meynard J.M., 2001b. Statistical methods for predicting responses to applied nitrogen and calculating optimal nitrogen rates. *Agron. J.*, **93**, 531-539.
- Makowski D. & Wallach D., 2002. It pays to base parameter estimation on a realistic description of model errors. *Agronomie*, **22**, 179-189.
- Makowski D. et al., 2006. Parameter estimation for crop models. In: Wallach D., Makowski D. & Jones J.W. *Working with dynamic crop models. Evaluation, analysis, parameterization, and applications*. Amsterdam, The Netherlands: Elsevier.
- Metropolis N. et al., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087-1092.
- Naud C., Makowski D. & Jeuffroy M.H., 2008. Is it useful to combine measurements taken during the growing season with a dynamic model to predict the nitrogen status of winter wheat. *Eur. J. Agron.*, **28**, 291-300.
- Ritchie J.T. & Otter S., 1984. *Description and performance of CERES-Wheat, a user oriented wheat yield model*. Temple, TX, USA: USDA-ARS-SR Grassland Soil and Water Research Laboratory.
- Ruget F., Brisson N., Delécolle R. & Faivre R., 2002. Sensitivity analysis of a crop simulation (STICS) in order to determine accuracy needed for parameters. *Agronomie*, **22**, 133-158.
- Ruget F., Stager S., Volaire F. & Lelievre F., 2009. Modelling tiller density, growth, and yield of Mediterranean perennial grasslands with STICS. *Crop Sci.*, **49**(6), 2379-2385.
- Sinclair T.R. & Seligman N.A., 2000. Criteria for publishing papers on crop modeling. *Field Crops Res.*, **68**, 165-172.
- Singh A.K., Tripathy R. & Chopra U.K., 2008. Evaluation of CERES-Wheat and CropSyst models for water-nitrogen interactions in wheat crop. *Agric. Water Manage.*, **95**, 776-786.
- Van Diepen C.A., Wolf J., Van Keulen H. & Rappoldt C., 1989. WOFOST: a simulation model of crop production. *Soil Use Manage.*, **5**, 16-24.
- Varella H., Guérif M. & Buis S., 2010. Global sensitivity analysis measures the quality of parameter estimation: the case of soil parameters and a crop model. *Environ. Modell. Softw.*, **25**, 310-319.
- Vrugt J., ter Braak C., Gupta H. & Robinson B., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling. *Stochastic Environ. Res. Risk Assess.*, **23**.
- Wallach D. et al., 2001. Parameter estimation for crop models: a new approach and application to a corn model. *Agron. J.*, **93**, 757-766.
- Wallach D., Makowski D. & Jones J.W., 2006. *Working with dynamics crop models. Evaluation, analysis, parametrization and applications*. Amsterdam, The Netherlands: Elsevier.
- Wallach D. et al., 2011. A package of parameter estimation methods and implementation for the STICS crop-soil model. *Environ. Modell. Softw.*, **26**, 386-394.
- Weir A.H., Bragg P.L., Porter J.R. & Rayner J.H., 1984. A winter wheat crop simulation model without water or nutrient limitations. *J. Agric. Sci.*, **102**, 371-382.
- Weiss A. & Wilhem W.W., 2006. The circuitous path to the comparison of simulated values from crop models with field observations. *J. Agric. Sci.*, **144**, 475-488.
- William J.R., Jones C.A., Kiniry J.R. & Spanel D.A., 1989. The EPIC crop growth model. *Trans. ASAE*, **32**, 497-511.
- Yuping M. et al., 2008. Monitoring winter wheat growth in North China by combining a crop model and remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.*, **10**, 426-437.
- Zadoks J.C., Chang T.T. & Konzak C.F., 1974. A decimal code for the growth stages of cereals. *Weed Res.*, **14**, 415-421.

(44 réf.)