Published in: Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009), Egmond aan Zee, The Netherlands, 10-12 March 2009. - Wageningen : Wageningen UR, 2009 - ISBN 978 (2009).

Parameter estimation software for crop models

D. Wallach, F. van Evert
INRA UMR AGIR
BP 52627
31326 Castanet Tolosan Cedex France
wallach@toulouse.inra.fr

Introduction

Parameter estimation is a major aspect of crop modelling. Together with the functional forms of the equations, it is a major determinant of prediction quality. It is also often very time consuming. Finally, it requires expertise quite different than that required for proposing model equations, fixing objectives of a modelling exercise or analysing the results. This suggests that it is of major interest to provide software to automate model parameter estimation. That is the purpose of the work reported here.

The software reported here is connected to the MODCOM modelling platform (Hillyer, Bolte, van Evert and Lamaker 2003), which is used in the SEAMLESS project. The platform includes a simulation motor based on the DEVS paradigm (Zeigler and Praehofer 2000). A major attraction of linking the software to MODCOM is that in this way it is not specific to a particular model. Rather, it is available to any model put on the MODCOM platform, or to individual components of a model.

Parameter estimation for crop models is generally based on a least squares approach, but the problem is complicated by the complexity of the data and the models, and also by the fact that there are two types of information about the parameters, namely information based on studies of the individual processes and information based on studies of the whole system. Different algorithms may be appropriate in different cases, and for complex data there is no standardized approach. Therefore the software includes several different estimation algorithms, and others can easily be added.

Methods

The parameter estimation algorithms are written in the R statistical computing language (R Development Core Team 2007), which is freely available. The MODCOM software is written in the C sharp language. Information is passed from one to the other using Microsoft COM technology. The R software requires the following information, which must be stored in a data base which is part of the software: the observed data, the name of the model to be run, the paths to the input files for the contexts of the data (for example climate, soil and management files for each context), the list of model parameters and indicators as to which are to be estimated and finally information related to the correspondence between the data and the model output. At each iteration the R routine sets the parameter values in MODCOM to the current values, executes the model for each context, retrieves the results, calculates the criterion to be minimized and determines the parameter values for the next iteration. The search algorithms are built into available R functions.

The different parameter estimation methods correspond to different hypotheses about model error (Makowski, Hillier, Wallach, Andrieu and Jeuffroy 2006): all errors are independent with the same normal distribution (ordinary least squares), different measured variables have different variances and all errors are independent (generalized least squares), different measured variables have different variances and errors for a given measurement variable are not independent (generalized least squares-like criterion), each context has the same variance-covariance matrix for model error (determinant criterion). In the generalized least squares case, the variances can be estimated from the data or fixed by the user. Also, the data can first be log transformed or not. We are also experimenting with Bayesian estimation procedures.

The R functions are separated into those that handle communication with the model and those that actually perform the search over the parameter space. One can then connect the software to alternative modelling platforms by changing only the communication routines. The software has thus also been adapted to the RECORD platform, which is also based on DEVS (Chabrier, Garcia, Martin-Clouaire, Quesnel and Raynal 2007).

Results and discussion

It is envisioned that the software be used in several different ways. First, it can be used to estimate parameters for individual model components, using data that concern a single component decoupled from the rest of the model. Second, it can be used to estimate parameters using the full model and multiple types of measurements. Finally, it can be used as a methodological tool to compare different estimation algorithms.

When a model does not agree well with data, the problem no doubt arises from both the way the equations describe the processes and from the values of the parameters. It is hoped that by providing a convenient and flexible set of parameter estimation algorithms, the software we propose will allow modellers to deal more simply and more efficiently with the parameter value aspect of the problem. It should then be possible to better evaluate the role of the process equations in model error.

References

Chabrier, P., Garcia, F., Martin-Clouaire, R., Quesnel, G., and Raynal, H. (2007), Toward a simulation modeling platform for studying cropping systems management: the Record project, presented at Farm Systems Design 2007, an international symposium on Methodologies for Integrated Analysis of Farm Production Systems,Catania,Sicily,Italy,10-12 September 2007.

Hillyer, C., Bolte, J., van Evert, F., and Lamaker, A. (2003), 2003. The ModCom modular simulation system. European Journal of Agronomy 18:333-343. *European Journal of Agronomy*, 18, 333-343.

Makowski, D., Hillier, J., Wallach, D., Andrieu, B., and Jeuffroy, M. H. (2006). Parameter estimation for crop models, in *Working with dynamic crop models*, D. Wallach, D. Makowski, and J. W. Jones (eds.), Amsterdam: Elsevier.

R Development Core Team . A language and Environment for Statistical Computing. 2007. Vienna, Austria, R Foundation for Statistical Computing.

Zeigler, B. P. and Praehofer, H. (2000), *Theory of modeling and simulation,* Academic Press.