FACCE-MACSUR

Data classification and criteria catalogue for data requirements

Task leader: Claas Nendel[1]* (P147)
Partner involved:
Technical University of Madrid (24)
German Weather Service (99)
Global Change Research Centre, Academy of Sciences of the Czech Republic (P 17)
Institut National de la Recherche Agronomique, INRA (P 175/188/169/196/206/201)
Scottish Agricultural College (P 47)
Norwegian Institute of Agricultural and Environmental Research, Bioforsk (P 128)
Aarhus University (P 189)
University of Bonn (P 115)
Swedish University of Agricultural Sciences and Lund University (P 163/94)
Institute of Soil Science and Plant Cultivation- Research Institute Pulawy (P 125)
Karlsruhe Institute of Technology (P 101)

[1] Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Straße 84, 15374 Müncheberg, Gemany

*nendel@zalf.de

Instrument:	Joint Programming Initiative
Topic:	Agriculture, Food Security, and Climate Change
Project:	Modelling European Agriculture with Climate Change for Food Security (FACCE-MACSUR)
Due date of deliverable:	month 6
Start date of project:	1 June 2012
Duration:	36 months
Work Package:	CropM 1
Deliverable lead partner:	Leibniz-Zentrum für Agrarlandschaftsforschung
Deliverable reference num.:	D-C1.2
Deliverable type:	Report
Confidential till:	—

Revision	Changes	Date
1.0	First Release	2013-04-20

Abstract

Data requirements for calibration and validation of agro-ecosystem models were elaborated and a classification scheme for the suitability of experimental data for model testing and improvement has been developed. The scheme enables to evaluate datasets and to classify datasets upon their quality to be used in crop modelling.

Table of Contents

Introduction

In crop modelling precise experimental agricultural datasets are essential for calibration and validation of models. Since agricultural datasets were recorded under most varying demands, its level of detail, quality of records, number of parameters considered as well as the number of spatial and temporal replicates varied enormously.

In general, datasets used for model calibration and validation comprise of data describing a) the initial soil conditions, b) the crop-specific management and c) the seasonal climate (Palosuo et al. 2011, Rötter et al. 2012). Additionally, the phenology of the crop, yields and nutrient contents from intermediate harvests, intra-seasonal soil conditions et cetera may be provided.

Requirements on the dataset strongly depend on the application: For model calibration requirements on the experimental data are much higher than for model validation. Beside the state variables of the crops themselves also observations of the boundary conditions for crop growth like weather conditions and soil water and nutrient status are important to test consistency of model simulations.

Thus, before applying certain datasets in crop modelling it is recommended to clarify the quality and to classify it.
The aim of Task C1.2 was to elaborate a minimum requirement for data sets for model testing as well as useful observations for calibration and validation and to develop a classification scheme by which the consistency and quality of agricultural datasets can be evaluated.

Methods

A classification scheme was developed to simplify the assessment of quality and level of detail of datasets. The scheme encompasses ten groups of input variables, namely:

- Cultivation
- Phenology
- Preceding crop
- Initial values
- Soil
- Site data
- Weather data

And the following groups of state variables to be compared with model outputs:

- Crop
- Soil
- Observations

Within these groups 65 potential variables were listed which may be describing the experimental agricultural datasets.

For each variable an estimation of their importance for crop modelling was carried out. According to the importance a weight (from 1 to 5) was assigned to each variable. For instance, in crop modelling the weather parameters Tmax and Tmin (the daily maximum and minimum temperature) are of great importance. Thus, the variables get the highest weight (5), whereas Tavg (daily average temperature) has lower predictive power and was assigned the weight 1. Additionally, the weighting points are modified individually to cover aspects like temporal and spatial resolution, representativeness, accuracy of methods of observations.

Next, 4 classes were defined to classify the datasets by quality: From low to high quality theses classes encompass the levels "bronce", "silver", "gold" and "platinum".

Accordingly, for each group of variables and each class a minimum amount of information was defined by a sum of weight points. This should ensure that data sets are well balanced and relevant deficits in one group of input variables are not masked by very detailed information of another group. Thus, by listing all variables given in an experimental dataset and multiplying these variables with the above-mentioned weights the dataset automatically becomes classified into one of the four quality classes.

In addition, a minimum dataset was defined. This hypothetical dataset describes the minimum amount of information that is needed in agricultural datasets to be useful for any modelling.

Results

On the basis of the weighting of the specific variables an Excel file was designed to facilitate the classification of datasets. In Fig. 1 the classification scheme is demonstrated including an example dataset. Further, the quality classes, its minimum sums of weight points and the classification of the sample dataset are shown.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Classification of datasets, CropM (MACSUR)**												
2			minimum data	depth (m)	no. of layers or observations	weight point	replicates	result weight point x (col. D and/ or col. E) *x col. F/3 (opt.)*	sub-block sum	minimum block sum bronce	minimum block sum silver	minimum block sum gold	minimum block sum platinum
3	cultivation	variety			0	1		0					
4		sowing	x		1	5		5					
5		harvest	x		1	5		5					
6		fertilization	x		1	5		5					
7		irrigation	x		0	5		0					
8		seed density			1	1		1					
9		yield	x		1	5		5					
10		tillage				1		0	21	21	22	23	23
11	phenology	emergence			0	3		0					
12		tillering/stem elongation			0	1		0					
13		ear emergence			0	3		0					
14		flowering	x		1	5		5					
15		yellow ripeness			1	3		3	8	8	11	11	14

Fig. 1: Screenshot of classification scheme for agricultural dataset

Discussion

The classification scheme was discussed intensively among WP1 and WP2 leaders. Its limits were identified in the subjectivity of the weight-points. The importance of specific input variables depends on their ecological significance but also on the model assumptions, the model architecture and the interested output variable/research question.

It was decided to test the scheme with collected data sets within MACSUR for model comparison and to further develop the scheme according to the experience of the project. Weighting points will be justified using literature describing variance and standard errors of the different state variables and measurement methods (e.g. Dahiya et al. 1984, Roth et al. 1992, Jacobsen & Schjøninng 1993, Wegehenkel 1998, Wendroth et al. 2001, Kersebaum et al. 2002, Giebel et al. 2006).

A preliminary version of the scheme has been uploaded to the MACSUR homepage; filename: macsur_data_criteria_V1.xls

This work will be finalised with a scientific paper analysing the accuracy and potential errors of different measurements and methods related to the sensitivity of different models and presenting an applicable scheme of data classification.

Title: Analysis and classification of data sets for calibration and validation of agro-ecosystem models.

Authors: K.C. Kersebaum, M. Bindi, C. Nendel, J.E. Olesen, M. Trnka, C. Kollas, T. Gaiser, C. Frühauf, F. Ruget + contributors to the discussion and analysis of this task.

Journal: Multi-disciplinary or agricultural journal (e.g. European Journal of Agronomy, Ecological Modelling)

Deadline: first draft by September 2013

Acknowledgements

We thank all the CropM partners for contributions to the discussion of the classification scheme.

This paper is a contribution to the FACCE MACSUR knowledge hub.

The work was funded by the Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz, Germany.

References

DAHIYA, I. S., K. C. KERSEBAUM & J. RICHTER (1984): Spatial variability of some nutrient constituents of an Alfisol from loess: I. Classical statistical analysis. Z. Pflanzenernähr. Bodenk. 147: 695 – 703.

GIEBEL, A., WENDROTH, O., REUTER, H.I., KERSEBAUM, K.C., SCHWARZ, J. (2006): How representatively can we sample soil mineral nitrogen? J. Plant Nutr. Soil Sci. 169: 52-59.

JACOBSEN, O.H., SCHJØNNING, P. (1993): Field evaluation of time domain reflectometry for soil water measurements. J. Hydrol. 151: 159-172.

KERSEBAUM, K.C., REUTER, H.I., LORENZ, K., WENDROTH, O. (2002): Modelling crop growth and nitrogen dynamics for advisory purposes regarding spatial variability. In: L. Ahuja, L. Ma & T. Howell (eds.): Agricultural system models in field research and technology transfer. Lewis Publishers, Boca Raton, 229 - 252.

PALOSUO, T., KERSEBAUM, K.C., ANGULO, C., HLAVINKA, P., MORIONDO, M., OLESEN, J.E., PATIL, R.H., RUGET, F., RUMBAUR, C., TAKÁČ, J., TRNKA, M., BINDI, M., ÇALDAĞ, B., EWERT, F., FERRISE, R., MIRSCHEL, W., ŞAYLAN, L., ŠIŠKA, B., RÖTTER, R. (2011) Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. European Journal of Agronomy, 35: 103- 114

RÖTTER R.P., T. PALOSUO, K.C. KERSEBAUM, C. ANGULO, M. BINDI, F. EWERT, R. FERRISE, P. HLAVINKA, M. MORIONDO, C. NENDEL, J.E. OLESEN, R.H. PATIL, H, F. RUGET, J. TAKÁČ, M. TRNKA (2012). Simulation of spring barley yield in different climatic zones of Northern and Central Europe: A comparison of nine crop models. Field Crops Res. 133: 23-36

ROTH, C.H., MALICKI, M.A., PLAGGE, R. (1992): Empirical evaluation of the relationship between soil dielectric constant and volumetric water content as the basis for calibration soil moisture measurements by TDR. J. Soil Sci. 43. 1-13.

WEGEHENKEL, M. (1998): Zum Einsatz von TRIME-TDR zur Messung der Bodenfeuchte auf leichten Sandstandorten. J. Plant Nutr. Soil Sci. 161, 577-582.

WENDROTH, O., JÜRSCHIK, P., KERSEBAUM, K.C., REUTER, H., VAN KESSEL, C., NIELSEN, D.R. (2001): Identifying, understanding, and describing spatial processes in agricultural landscapes - four case studies. Soil Tillage Research 58, 113 - 128