Europ. J. Agronomy 77 (2016) 209–222

Contents lists available at ScienceDirect

European Journal of Agronomy

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

Interoperability of agronomic long term experiment databases and crop model intercomparison: the Italian experience

Fabrizio Ginaldi[a,*], Marco Bindi[b], Anna Dalla Marta[b], Roberto Ferrise[b], Simone Orlandini[b,c], Francesco Danuso[d]

[a] Council for Agricultural Research and Economics, CREA-AA, Via di Corticella 133, 40128 Bologna, Italy
[b] University of Florence, Department of Agri-food Production and Environmental Sciences (DISPAA), Piazzale delle Cascine, 18, 50144 Firenze, Italy
[c] Fondazione per il Clima e la Sostenibilità, Via G. Caproni 8, 50145 Firenze, Italy
[d] University of Udine, Department of Agricultural and Environmental Sciences (DISA), Via delle Scienze 206, 33100 Udine, Italy

ARTICLE INFO

Article history:
Received 1 October 2015
Received in revised form 8 February 2016
Accepted 23 February 2016
Available online 10 March 2016

Keywords:
Data structure
Database
Crop model
Interoperability
Long term field experiments

ABSTRACT

The IC-FAR national project (Linking long term observatories with crop system modelling for better understanding of climate change impact, and adaptation strategies for Italian cropping systems) initiated in 2013 with the primary aim of implementing data from 16 long term Italian agronomic experiments in a common, interoperable structure. The building of a common database (DB) structure demands a harmonization process aimed at standardising concepts, language and data in order to make them clear, and has to produce a well-documented and easily available tool for the whole scientific community. The Agricultural Model Intercomparison and Improvement Project (AgMIP) has made a great effort in this sense, improving the vocabulary developed by the International Consortium for Agricultural Systems Applications (ICASA) and defining harmonization procedures. Nowadays, these ones have also to be addressed to facilitate the extraction of input files for crop model simulations. Substantially, two alternative directions can be pursued: adapting data to models, building a standard storage structure and using translators that convert DB information to model input files; or adapting models to data, using the same storage structure for feeding modelling solutions constituted by combining model components, re-implemented in the same model platform.

The ICFAR information management system simplifies data entry, improves model input extraction (implementing System Dynamics ontology), and satisfies both the paradigms. This has required the development of different software tools: ICFAR-DB for data entry and storage; a model input extractor for feeding the crop models (MoLInEx); SEMoLa platform for building modelling solutions and performing via scripts the model intercomparison. The use of the standard AgMIP/ICASA nomenclature in the ICFAR-DB and the opportunity to create files with MoLInex for feeding AgMIP model translators allow full system interoperability.

Software and/or data availability

ICFAR database template is an Excel ™ xlsx workbook, available upon request from the corresponding author or the IC-FAR project coordinator Pier Paolo Roggero.

MoLInEx data extractor is an executable software running under Windows OS. MoLInEx extracts data from the standard ICFAR-DB version 2.0 and is freely available from the corresponding author.

DBupdater executable performs an automatic update of the Management sheet in ICFAR-DB by web downloading the latest version of practices and applications.

All the DBs and software described herein are stored in the on-line repository available on http://www.icfar.it/.

1. Introduction

The national project IC-FAR, granted by the Italian Ministry of Education, Universities and Research, initiated in 2013 with the aim of implementing data from 16 long term Italian agronomic experiments (LTEs) in a common, interoperable structure and to assess the reliability of different cropping system models over a wide range of

* Corresponding author. Tel.: +39 0425 360113; fax: +39 0425 34691.
E-mail address: fabrizio.ginaldi@crea.gov.it (F. Ginaldi).

http://dx.doi.org/10.1016/j.eja.2016.02.007

Mediterranean environments and cropping systems. The selected models are intended to be used for scenario and uncertainty analyses with respect to near-future climate change. The LTEs are located in seven Italian sites: Torino, Padua, Bologna, Ancona, Pisa, Perugia and Foggia. The oldest LTEs have been operating since the 1960s, at a latitude range between 41° and 45°N. The network of LTE involves many agricultural management practices concerning crop, cultivar, soil labour, fertilisation, irrigation, and other. Treatments are repeated, year by year, on the same soil plots (in this issue Baldoni et al., 2016a,b; Bonciarelli et al., 2016; Mazzoncini et al., 2016; Pituello et al., 2016; Seddaiu et al., 2016; Ventrella and Vitti, 2016; Vitali et al., 2016; Zavattaro and Grignani, 2016).

Despite the huge amount of information obtained, at least as far as Italian trials are concerned, very little effort has been made to put data into a common format, thus making them available for further studies to be performed by the scientific community. The most recent one, which led to the development of a relational meta-database for data management of the LTEs (ClimAgriLT; Zuliani et al., 2003), was discontinued, due to the lack of interest shown by LTE managers towards adopting and populating this database. This was probably due to two main reasons: the type of support adopted (DBMS based on MS-Access, difficult to use for many researches who often prefer using spreadsheets) and the lack of consistent and satisfactory rules for sharing data obtained from the different research units.

Therefore, on the basis of previous Italian experiences and results, IC-FAR project partners decided to create a new standard structure for LTE databases to store data, spreadsheet based, with the requisite of ensuring compatibility with other international database platforms (AgMIP, http://www.agmip.org/, Rosenzweig et al., 2013; MACSUR, http://macsur.eu/, Rötter et al., 2013). Therefore, not only does ICFAR-DB try to implement previous ClimAgri project experience, but also to integrate it with that deriving from the MACSUR project and GRA Croplands Research Group on Agricultural Greenhouse Gases (Franzluebbers et al., 2015, http://globalresearchalliance.org/). The goals for such a new structure were: (i) to make it easy for LTE managers to record data in a standard database structure; (ii) to facilitate and speed up the extraction of input files to feed cropping systems models; (iii) to improve the current scarce utilisation of Italian LTE data to develop, calibrate and validate cropping system models, even in international contexts.

The use of simulation models to describe cropping systems is usually considered mandatory to integrate the complex information derived from LTE experiments and to realistically predict behaviour and productivity of an agricultural system, facing the current rapid climate changes (Argent, 2004a,b; Rötter et al., 2011). This task can be pursued by evaluating existing models (by a model intercomparison process, for a reference protocol see Bellocchi et al., 2014) or by developing new models in a modular fashion. At present, in the IC-FAR project, every research unit (RU) has experience in using different cropping system models: DSSAT (Jones et al., 2003) at Bari and Sassari, EPIC (Williams, 1995) at Sassari, CropSyst (Stöckle et al., 2003) at Firenze, SALUS (Basso et al., 2006) at Potenza, CSS (Cropping System Simulator, Danuso et al., 1999) at Udine. The Udine RU is also involved in the development and maintenance of CSS, based on the SEMoLa modelling platform (semola.uniud.it, Danuso, 1992; Danuso and Rocca, 2014).

This paper initially highlights the difficulties that ICFAR project partners had to handle: (i) to build a common standard structure in which to store data; (ii) to define the language and procedures to use in the data harmonization process to generate crop model input files; (iii) to ensure both model interoperability, using the same experimental input dataset for feeding several crop models, and project interoperability to make datasets immediately available for use in the context of other international initiatives on field experimental data sharing and model intercomparison; (iv) to identify the best course of action to pursue for performing model simulations and to produce the most reliable results.

After a state-of-the-art presentation on these topics, with the addition of some further considerations, the ICFAR project informative system is presented.

1.1. Concepts and language harmonization

Harmonization is a process aimed at standardising concepts, language and data in order to make them well documented, understandable and available to the whole scientific community.

Knowledge transfer among research groups and models requires the definition of a standard vocabulary and an ontological structuring of information (separating them into classes with pre-defined properties).

Common standard vocabulary is a pre-requisite which enables transferring concepts, linked to clearly defined words. In this field, hierarchies can be used to organise terms but, moreover, it is important to define specific labels (also said names, ID, acronyms) used in information systems dealing with quantitative entities (Schentz et al., 2015). In order to enhance an efficient and effective use of model ensembles, the AgMIP project is performing a data harmonization process that facilitates data storage and access, interoperability and simulation interpretation. The use of model ensembles has been proposed for a better characterization of uncertainty associated with model outputs and a more accurate estimate of crop responses (Asseng et al., 2013; Bassu et al., 2014). In the agronomy field, the standardisation process, which was begun by the ICASA consortium in 1995, and is still in progress, has demanded a basic definition of a common language in the field of cropping systems (Hunt et al., 2001; White et al., 2013).

Nonetheless, a further step forward should be taken considering the development of an ontology that may permit an easier management of cropping system knowledge for modelling. Ontology is considered as an explicit formal specification of how to represent objects, concepts and their relationships, existing in some areas of interest (Gruber, 1992). Ontology facilitates this task defining, *a priori*, the properties of each ontological class, so outlining how to manage information pertaining to the same class (i.e., the mineralization coefficient and the light extinction coefficient, even being different, pertain to the same class of "parameters" so they share some common properties and can be managed as all other parameters)

1.2. Harmonization procedures

Several publications recently identified a number of substantial concerns associated with data sources and methods used in studies on climate and land use change involving crop simulation models (van Ittersum et al., 2013; Van Wart et al., 2013). These concerns include: (i) poor quality of weather and soil data; (ii) unrealistic assumptions about the cropping-system context; (iii) poorly calibrated crop simulation models; and (iv) lack of transparency about underpinning assumptions and methods for re-building missing data, on model input dataset quality, and adopted calibration procedures (Grassini et al., 2015). Surprisingly, despite the wide use of crop simulation models, there are no published guidelines about standard sources and quality of data input for weather, soil, actual yields, and cropping-system context, or requirements for calibration of crop models used in such studies.

Hence, not only is there a need to harmonize languages and data but also to harmonize the procedure in performing model intercomparison exercises.

Both model intercomparison and potentially new model development require the definition of a common input dataset, whose