ECOFI: a generic agronomic database to facilitate analysis and crop modelling

S. Auzoux [1] – *L. Rouan* [2] – *R. Loison* [1] – *J.-F. Martiné* [1]

[1] UR AÏDA, CIRAD, Av. Agropolis, 34398 Montpellier, France, sandrine.auzoux@cirad.fr
[2] UMR AGAP, CIRAD, Av. Agropolis, 34398 Montpellier, France

Introduction

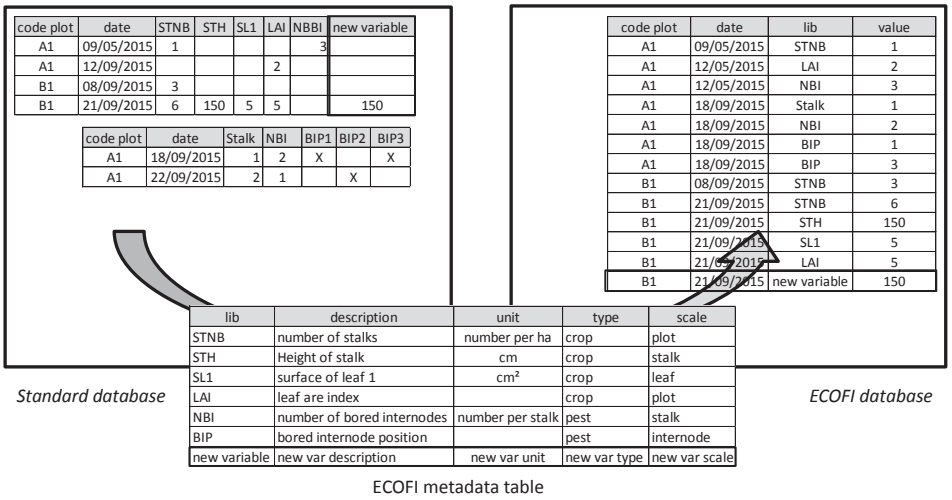
Studies of genotype x environment x management interactions and agroecology commonly use complex models such as Crop Simulation Models (CSM). Each model requires reliable minimum datasets (MD) for its successful implementation (Grassini et al., 2015; Nix, 1983). These MD are collected separately and can be multi-scale, multi-species, multi-disciplinary (agronomy, entomology, phytopathology, weed science, etc.) thereby making their use difficult in modelling. Furthermore, all variables are not measured simultaneously leading to the occurrence of many empty cells. All these problems can be solved using database technology (Hunt et al., 2001). This paper describes how the generic agronomic database ECOFI was implemented.

Materials and Methods

The database schema of ECOFI was built from the content of many CSM input files and field experiment datasets collected through studies in agronomy, entomology and phytopathology. We observed that although they are organized differently, most of the resulting agronomic databases shared the same measurements (yield, leaf area index, biomass, insect incidence, etc.) and a few similar tables corresponding to the minimum dataset (weather, soil, crop, and management data). Based on this analysis, we have designed the structure of ECOFI. It's divided into two parts: the first describes environmental conditions while the second describes all the possible cropping practices and agronomic measurements. It was implemented using the open source object-relational database management system PostgreSQL (©1996-2015 The PostgreSQL Global Development Group).

Results and Discussion

In standard databases, each additional observed variable implies to update the existing database schema or to create a new table. In ECOFI, we can add a new variable simply by adding a new record in one table (Figure 1). All variable labels are stored in a metadata table including the units of measurement, the type of variable observed and the scale of observation.



code plot	date	STNB	STH	SL1	LAI	NBBI	new variable
A1	09/05/2015	1				3	
A1	12/09/2015				2		
B1	08/09/2015	3					
B1	21/09/2015	6	150	5	5		150

code plot	date	Stalk	NBI	BIP1	BIP2	BIP3
A1	18/09/2015	1	2	X		X
A1	22/09/2015	2	1		X	

code plot	date	lib	value
A1	09/05/2015	STNB	1
A1	12/05/2015	LAI	2
A1	12/05/2015	NBI	3
A1	18/09/2015	Stalk	1
A1	18/09/2015	NBI	2
A1	18/09/2015	BIP	1
A1	18/09/2015	BIP	3
B1	08/09/2015	STNB	3
B1	21/09/2015	STNB	6
B1	21/09/2015	STH	150
B1	21/09/2015	SL1	5
B1	21/09/2015	LAI	5
B1	21/09/2015	new variable	150

lib	description	unit	type	scale
STNB	number of stalks	number per ha	crop	plot
STH	Height of stalk	cm	crop	stalk
SL1	surface of leaf 1	cm²	crop	leaf
LAI	leaf are index		crop	plot
NBI	number of bored internodes	number per stalk	pest	stalk
BIP	bored internode position		pest	internode
new variable	new var description	new var unit	new var type	new var scale

Figure 1. Standard database observation table versus ECOFI observation table using a metadata table

The technology of metadata is relevant as it allows the addition of as many data items as desired without changing the database schema. Another significant advantage is that it minimizes the number of tables, columns and empty cells. This makes it easier to export and manage the data. It also improves database query performance. ECOFI is available on a server and can also be used in disconnected mode when used with slow or no internet connection. ECOFI already has a wide application in pest management, plant disease and ecophysiological experiments on sugarcane, cotton and sorghum in Africa and Central America.

Conclusions

ECOFI is a performant optimized database that improves analysis and facilitates access to data for CSM. Genericness of database schema of ECOFI can allow intercomparison of CSM (AgMIP) that require the same datasets with no common data structure. It could also integrate other scales such as the gene or the landscape.

References

Nix, H. A. (1983). In: Proceedings of the International Symposium on Minimum Data Sets for Agrotechnology, 181-188.

Hunt, L.A., J.W. White, G. Hoogenboom (2001). Agricultural Systems, 70 (2–3): 477-492

Grassini, P., LG.J. van Bussel, J. Van Wart et al., (2015). Field Crops Research, 177: 49-63.