

## **REMS – a research experiment management system to support the maintenance and development of crop growth and development models.**

**Greg McLean**<sup>1,3</sup> and Alastair Doherty<sup>2</sup>

<sup>1</sup> Queensland Department of Primary Industries, PO Box 102 Toowoomba, Qld 4350

<sup>2</sup> School of Land and Food Sciences, The University of Queensland, Brisbane, Qld 4072

<sup>3</sup> Author for correspondence; [Greg.McLean@dpi.qld.gov.au](mailto:Greg.McLean@dpi.qld.gov.au)

### **Abstract**

The successful development and maintenance of crop growth and development models requires information from detailed physiological experiments conducted under varying environments and conditions. These datasets are usually collected by many different researchers, each using their own style of data storage, ranging from databases and spreadsheets, through to text files and field books. The Research Experiment Management System (REMS) was written to assist model developers by providing a common storage system for datasets. This system had to be easy to add to, provide graphical data display and with flexible output files designed as crop model inputs. Information in REMS is stored in three sections: general information (e.g. region, site, soil), experiment information (e.g. design, planting, irrigation), and data (e.g. crop, soil, meteorological). Program Plug-ins are used to gather experimental information that does not easily fit into these sections or that requires pre-processing. User written output templates are combined with REMS information to format output files that are suitable for crop models or data analysis. REMS has been successfully used with APSIM and the report package APSIMReport to provide graphs of observed and predicted trait data as an aid to model development and validation.

### **Media summary**

Development and maintenance of crop models will be easier because the REMS software package will store information from diverse experiments in an easy to use system. Flexible outputs formatted to suit crop model input files, and links to report packages means that the model developer can easily test the model against observed data.

### **Keywords**

Data storage, crop models, APSIM

### **Introduction**

Results from detailed experiments are the basic tools that crop model designers use in the development and validation of their models. These data need to comprehensively describe the environment, management practices and measured results of the experiment. The results may come from experiments conducted in various regions or countries by different researchers and stored in many different ways. To be able to use these tools to test and develop models, the data have to be stored uniformly, and be accessible to formulate model input files, and observed and predicted results.

REMS is broadly based on SugarBag(Laredo and Prestwidge 2003) and was developed as a system suitable for small workgroups. The aim during development was to keep everything as simple and easy to use as possible, while still providing the power and flexibility required to make it a useful program. To achieve this, effort was put into (i) making the data input simple, with no overall data dictionary; (ii) providing graphical views of the data as an aid in data validation; (iii) developing 'plug-ins' to pre-process and store data that does not fit the schema; and (iv) an output template procedure so that users can define the structure of their output files.

### **Materials and methods**

## The program

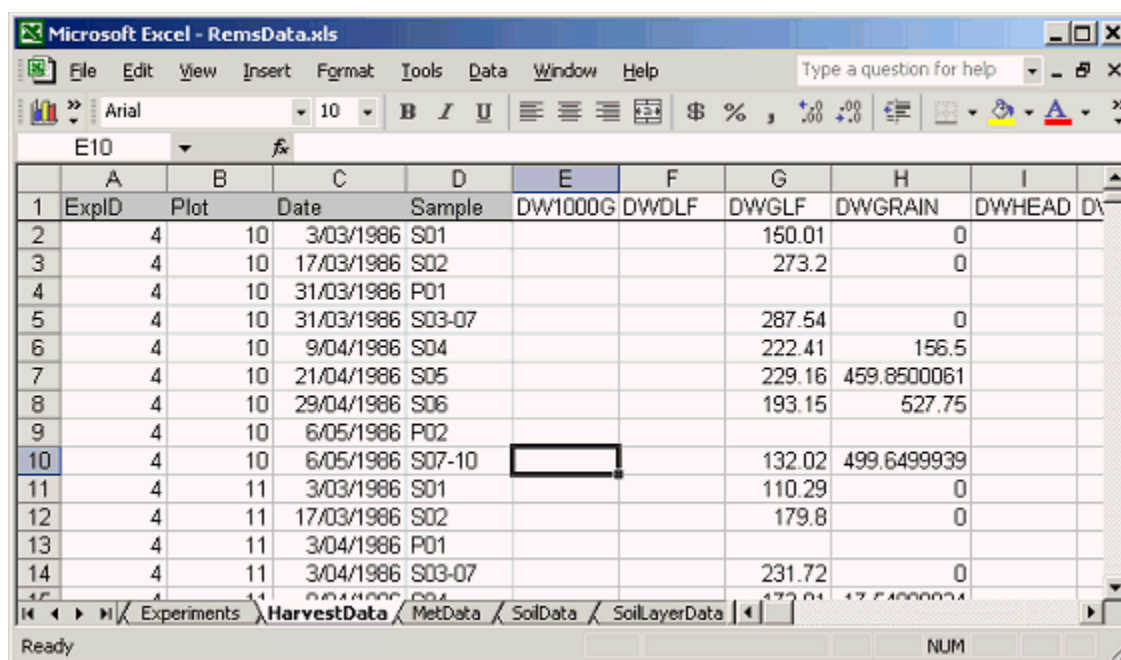
REMS is written in Borland C++ using ADO (Active Data Objects) to link to Microsoft Access files (.mdb). All data is stored in a fully relational database structure, employing referential integrity. The object orientated front-end, uses tree controls to provide intuitive access to the different types of stored information.

Three data classes are used: *Information* about regions, sites, soils, experimentalists etc.; *Experiment* description, including design and operations; and *Data* including soil, weather and plot data collected during the experiment.

## Input

Previous examples of data storage systems used data input languages in which information is set up in particular columns or with a specific prefix. (Van Evert *et al.*, 1999, Hunt *et al.*, 2001). It was observed that most experimentalists use spreadsheets to collect their data, so REMS was developed to use formatted spreadsheets for input experimental conditions and data. There are three spreadsheets, reflecting the system's data classes (Information, Experiment and Data)(Figure 1). Data can also be entered or modified within the system's user interface.

These spreadsheets have fixed pages for sub-classes, and fixed columns for compulsory data. On pages such as the HarvestData page, the user can enter columns of data, heading each column with the data name of their choice.



	A	B	C	D	E	F	G	H	I	J
1	ExpID	Plot	Date	Sample	DW1000G	DWDLF	DWGLF	DWGRAIN	DWHEAD	DW...
2	4	10	3/03/1986	S01			150.01	0		
3	4	10	17/03/1986	S02			273.2	0		
4	4	10	31/03/1986	P01						
5	4	10	31/03/1986	S03-07			287.54	0		
6	4	10	9/04/1986	S04			222.41	156.5		
7	4	10	21/04/1986	S05			229.16	459.8500061		
8	4	10	29/04/1986	S06			193.15	527.75		
9	4	10	6/05/1986	P02						
10	4	10	6/05/1986	S07-10			132.02	499.6499939		
11	4	11	3/03/1986	S01			110.29	0		
12	4	11	17/03/1986	S02			179.8	0		
13	4	11	3/04/1986	P01						
14	4	11	3/04/1986	S03-07			231.72	0		

Figure 1. Example data input spreadsheet

The data dictionary is left to the user and no system wide dictionary is employed. Users have flexibility in naming their variables and traits. A meta-database can be used if it is necessary to query a number of different REMS systems with dissimilar trait names.

The spreadsheet data is loaded in a hierarchical order, so it possible to validate the data, using various methods, as it is loaded into the REMS system. All of the names used to associate data are cross-checked with entries from the user defined data dictionary to maintain referential integrity.

## Views

Data views for each data class (ie. Information, Experiments and Data) are provided. (Figure 2) Each view provides graphs of the information, which can be used to check the integrity of the entered data, by identifying excessive variation in the data.

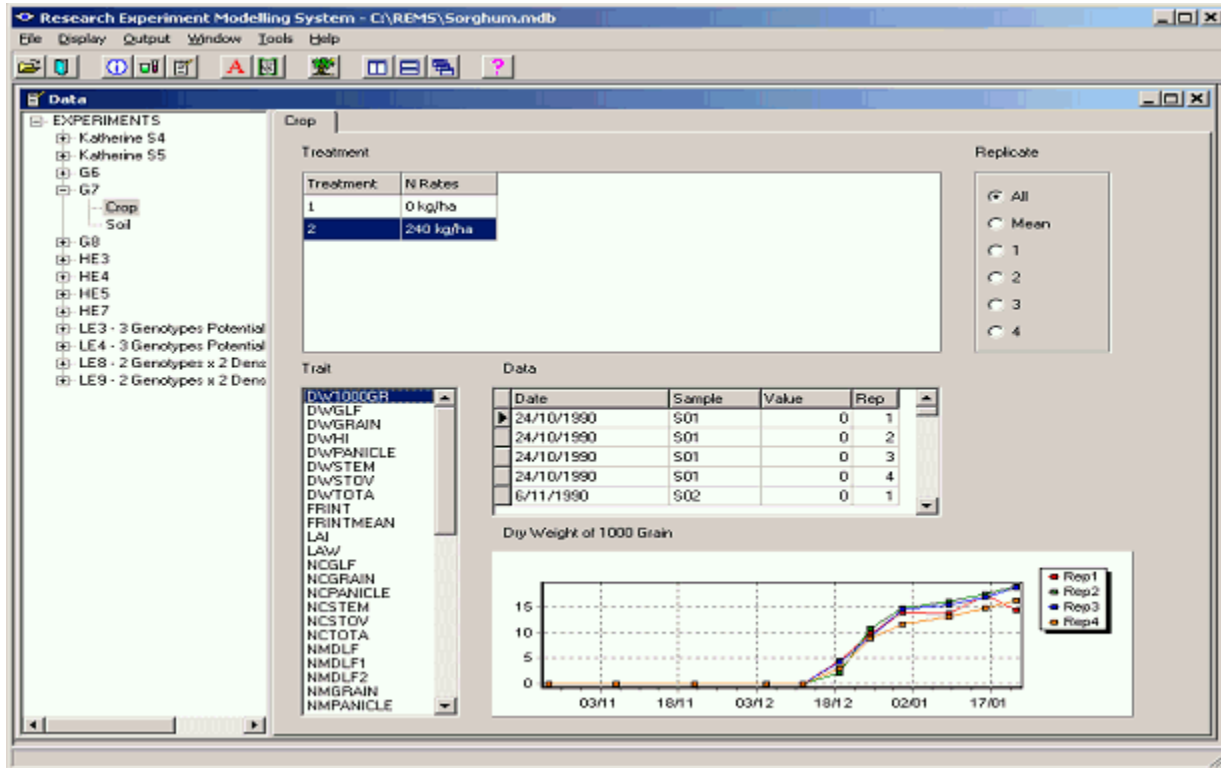


Figure 2. Data Views

## Plug-ins

To maintain the simplicity of the program and the data schema, plug-in programs are used to store and pre-process data that does not readily fit into the system design. For example, some users may wish to include sub-plot data or individual leaf area data (Figure 3). Instead of trying to allow for every data type in the REMS program, plug-in programs can be developed by the programming team and distributed for use.



Figure 2. Leaf Area Plug-in

### Output

Output templates are used to provide the flexibility needed to construct files suitable to run different crop models. The template language uses the information from REMS and a user designed template file to set up input files. Fig 4. shows an example template file used to write files for the APSIM crop simulation model(Keating *et al.*, 2003).

```
#file Session$.par

[Treatment$.Screen.parameters]
screen_output=on

[Treatment$.clock.parameters]
start_date = StartDate$
end_date = EndDate$

[Treatment$.operatns.parameters]
PlantingYear$ PlantingDay$ Crop$ sow plants = Population$ , sowing_depth = PlantingDepth$,
row_spacing = RowSpace$, cultivar = Cultivar$

SoilNitYear$ SoilNitDay$ soilwat2 set sw = Isw$
SoilNitYear$ SoilNitDay$ soiln2 set no3ppm = Ino3ppm$

#for_each Fertilize
```

```
FertYear$ FertDay$ fertiliz apply amount = FertNAmt$ (Units$), type = Urea_N, depth = FertDepth$ (mm)
#endfor

#endfile
```

#### Figure 4. Model template file

##### *The future*

Included in future development plans are better international support, improved reports and screens, including a summary report on all experiments and a link to web-based GIS system.

##### **Discussion**

The benefits of using the REMS system have become evident to the APSRU research group in several forms. The data uniformity that the REMS system provides has allowed easy extraction of often complex data to a usable and examinable form. Once the data is stored in REMS the experimental data is exercised regularly rather than being stored in a filing cabinet.

The referential integrity that REMS supplies is also a key factor in its reliability as a validation package. Data stored in spreadsheets, which is how the bulk of experimental data is collected, may contain references to misspelt experiment names or experiments that do not exist, which makes it very hard or impossible to extract data. The REMS system has various methods for checking and verifying data so that they can be as reliable as possible.

REMS has allowed scientists who are not entirely familiar with APSIM input files to quickly create the necessary run files and use their data to validate various APSIM modules. Model validation is the real strength of REMS, linked with APSIMReport, graphical views of observed vs. model predicted data can be easily produced and updated. The system is being used by many APSRU model developers and in numerous projects. Validation datasets are stored in REMS and used in APSIM software quality control.

##### **References**

Keating, B. A., Carberry P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N. G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J. P., Silburn, M., Wang, E., Brown, S., Bristow, K. L., Asseng, S., Chapman, S., McCown, R. L., Freebairn, D. M., and Smith, C. J. 2003. An overview of APSIM, a model designed for farming systems. *European Journal of Agronomy* **18** 267-288.

Laredo, L. A., Prestwidge, D. B. 2003. Sugarbag A Database System for Sugarcane Crop Growth, Climate, Soils and Management Data. CRC Sugar Occasional Publication - Brisbane, June 2003. 24.

Van Evert, F. K., Spaans, E. J. A., Kreiger, S. D., Carlis, J. V., Baker, J. M. 1999. A database for agro-ecological research data 1: data model. *Agron. J.* 91, 54-62.

Hunt, L. A., White, J. W., Hoogenboom, G. 2001. Agronomic data: advances in documentation and protocols for exchange and use. *Agricultural Systems* **70** 477-492.