Designing decision support systems for factorial analyses

N.I. Herrmann [a], E.J. Zurcher [a]

[a] *CSIRO, Canberra ACT*

Email: neville.herrmann@csiro.au

Abstract: Simulation modelling tools aim to give insights into system performance by simulating the system over a time period and then generating a set of results. When using a simulation modelling tool for decision support it is frequently necessary to process a number of simulations or scenarios in an experiment which can generate significant amounts of data. Constructing the scenarios and then examining their results frequently involves significant work to manage the input data and the result sets. It is not uncommon to do large amounts of post processing of the result data. Systems can be complex and factors contributing to their performance may be difficult to discover without a well designed reporting or statistical analysis tool.

Our solution to improving the understanding of systems by the user has been to integrate the experiment design, model processing and reporting of results within the decision support system. The integrated process in CSIRO's GRAZPLAN DSS tools is called an Analysis. By building a scenario and then applying variations to form a factorial experiment it is possible to provide a multidimensional solution space that can demonstrate which factors contribute to certain system behaviour. In the AusFarm and GrassGro DSS tools the Analysis process is fully supported by a multidimensional reporting system that combines results from all the scenarios within the experiment and displays comprehensive reports in a format that mostly eliminates the need for manual post-processing by the user.

Agricultural industry specialists, who are using GrassGro and AusFarm, have responded favourably to the effectiveness of this approach. Analysis reports show economic results and risks along with detail about the biophysical system for all the dimensions in the experiment. Users quickly develop a deeper understanding of the system being modelled. Solutions to management scenarios are being answered without the need for days or weeks of painstaking data management and post processing of outputs.

Keywords: *Decision Support System, DSS, Simulation Modelling*

1. INTRODUCTION

Decision support system software can model complex systems efficiently giving users deeper insights into their systems. To gain the most from many DSS tools, the results from modelled scenarios often need to be post-processed to provide transformed, aggregated and comparative data sets. The problem domain and the user determine the requirements for the information that should be summarised into the final output. Because of this it is often difficult for the designer of the DSS to provide a suitable reporting or summarising facility that fits all situations. This often leads to error-prone, time-expensive summarisation of the output data using spreadsheets or statistical packages.

In this paper we will demonstrate an approach to DSS tool design that aims to more completely inform users about the behaviour of their proposed scenarios while decreasing the effort required. The GrazPlan DSS tool GrassGro (Moore et al., 1997), developed by CSIRO, will be used to demonstrate that a more integrated approach to scenario exploration can bring DSS software tools from a single dimensional study into a multidimensional world. This brings DSS systems closer to the goal of answering management questions. GrassGro version 3 has been able to achieve this while improving the usability of the original design that previously only operated on a single simulated scenario at a time.

GrassGro version 3 has been adopted by agricultural industry advisory specialists and is providing quick turn-around times with efficient experiment design, improved reporting and little need for exhaustive post-processing of data. The ability of advisors and their clients to gain more insight into their management issues has been greatly improved.

2. EXISTING SYSTEMS

Since its release to the agricultural industry in 1997 (Donnelly et al., 2002), GrassGro has proved that it can reliably model an agricultural sheep or cattle grazing enterprise (Donnelly et al., 1998; Perillat et al., 2004). The original workflow process model for GrassGro followed along the same lines as many DSS systems of the past and present. This process focuses on building a representative farming system, entering initial state values and then executing the model over a time course. A data set is generated that shows the behaviour of the system and then output data can be presented in chart or table form in the software. The test scenario can then be modified and a new output set produced. Other DSS tools such as APSIM (Keating et al., 2003), GrazFeed (Freer et al., 1997), AusFarm (Moore et al., 2007) and WaterCast (Cook et al., 2009) are examples of systems that have followed this workflow model. Using this type of process the internal model can be validated and checked to determine how well it models the real world system.

This traditional workflow process model becomes unwieldy as problem complexity increases. Each of the previously mentioned systems has many initial values and output values. To view the generated output it is necessary to choose a set of variables and then display them in charts or tables in the DSS tool or post-process them using another software package. Many separate charts or tables may be required to get sufficient insight into the behaviour of the model with the test scenario. When a large set of scenarios needs to be considered to analyse a problem, the suites of input values and output results becomes very large, leading to many difficulties in modifying and controlling the input and output files. For example an agricultural advisor may be modelling the effects of seven different stocking rates on two types of pasture with two different shearing times. This generates 28 sets of output data. Managing the display of this output can be tedious and difficulty increased when this data is taken to post processing tools. After examination the execution and summarising may be repeated.

With the previous version of GrassGro, processing enough scenarios to demonstrate the range of management options that a typical landholder might consider usually meant a large investment of time and meticulous work. The post-processing of result sets normally took place using spreadsheets or statistical packages. With GrassGro having proved its modelling capability, we sought to improve the system to allow more efficient experiment design and reporting of results. The process of building sets of scenarios, executing them and then processing the suite of results was a very good candidate for automation. This would decrease areas where human errors could be introduced and minimise the user fatigue that was limiting the desire for using the system.

3. DESIGN OF THE NEW SYSTEM

During the investigation for a new system it was proposed that it was useful to think of a consultation with a traditional DSS as testing a single scenario. A scenario could be thought of as a plausible alternative path that the system could follow (Mahmoud et al., 2009). A software system that could test many scenarios could provide sensitivity analysis and it would also be possible to illustrate more clearly the interaction between the sub models in the system. The objective was to encapsulate the ‘scenario analysis’ into a single consultation.

The users of GrassGro were typically attempting to solve management questions. The definition of what a management question became the key to defining a new system. Management questions are what a producer would raise in relation to a particular scenario. The GrassGro training material instructed users how to determine answers to common questions such as; ‘*What stocking rate should I run?’* and *‘What is the best time to lamb for a prime lamb enterprise?’*. A generalisation of this was then called an Issue. A new terminology (Table 1) was introduced to capture key concepts in the consultation of GrassGro.

GrassGro Farm System	Describes the weather, soil, pasture, stock, and management rules of a certain enterprise type that could be tested over a time span for reliable representation of a real world system. This is closely related to the ‘simulation’ used in the previous version of GrassGro. A single scenario.
Management Question	The industry or environmental question being examined. For example: *What is the optimum stocking rate? What is the best shearing date?*
Issue	The framework for examining a Management Question within GrassGro. Defines the input constraints for applying variations to a *Tested Farm System,* and a design for a base report.
Analysis	Combining a Farm System and GrassGro Issue allows the response of the Tested Farm System to be tested over a range of conditions. Responses are summarised and can be compared directly in a specially designed report.

Table 1. Terminology used for the new GrassGro system.

The Use Case Analysis for the new software system (Figure 1) shows the three main roles and their interactions with the system.

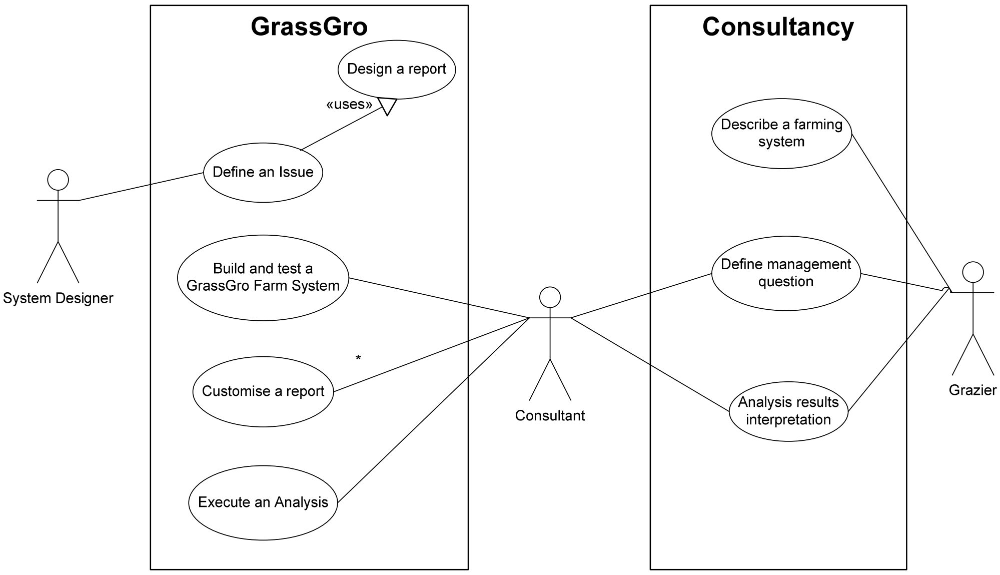


Figure 1. Use Case diagram for the GrassGro DS tool

The new design included making use of various libraries for users to store and reuse much of their previous work. Through using templates of reports and Issues along with components of farming systems it would be possible to reduce duplication, effort and improve the sharing of work between users.

The intention with a GrassGro Analysis was to allow a Farm System to be tested over a range of scenarios. Each tested scenario in an Analysis is called a treatment. Typically an Analysis would be two or three dimensional. Each dimension represents a factor that will be modified. Each of the factors can be represented by one of a number of factor values.

4. IMPLEMENTATION

The starting point for building a GrassGro Analysis is the definition of the Farm System. By decomposing this object into a number of smaller discrete component parts it became possible to use these components as iterated values or factor values in an Analysis. The component parts of a Farm System are not typically simple scalar values but larger objects composed of interdependent values. The disaggregation of a Farm System into these components required care so that dependencies between them were minimised. GrassGro can store these Farm System components in a Component library for reuse.

A vital part of the Analysis process is testing the Farm System for acceptability (see Figure 2). This Farm System must be one that is sufficiently representative of the real world system. To support the validation of the Farm System it is important to do Acceptability testing by simulating the Farm System over a time period. GrassGro provides detailed reports to assist in this step of the process. The user can examine in detail the biophysical system and the economic results in a single Acceptability report. Once the user thinks the Farm System is a sound representation they must explicitly tag the Farm System as being tested. This formalises the validation of the farming system and provides the basis for confidence in further scenario testing in the Analysis.

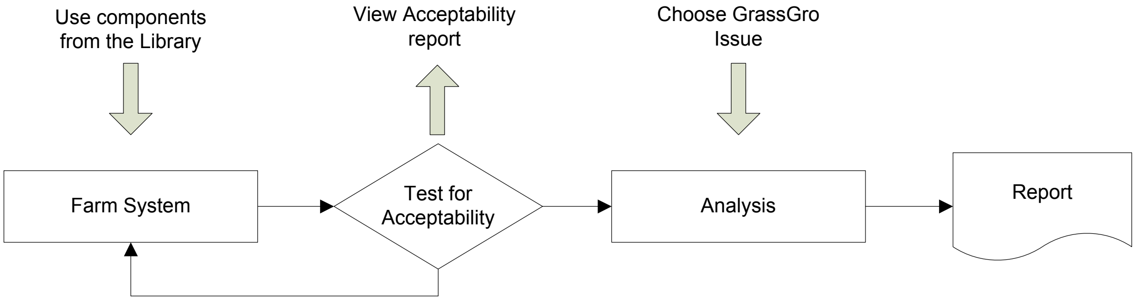


Figure 2. The process from building a Farm System through to generating an Analysis report.

In GrassGro, the Analysis is constructed using a GrassGro Issue, which is a template for a simulation experiment. Each Issue contains a definition of the Farm System component types that will be substituted in the simulation. These are the Factors which can have a number of Factor values. For instance a two-dimensional Analysis will have two Factors, each with a number of values. The Issue also contains a number of report templates that are matched to the type of Analysis. For example a stocking rate Analysis is best served by reports that show the effect of varying stocking rates. The numbers of dimensions that can be used for the Analysis are not limited, but as dimensionality increases so does the difficulty of interpretation of the results. The reporting system is built to dynamically handle the variation in numbers of dimensions and numbers of Factor values. This means that the user has no need to do any redesign of the report if the dimensionality of the Analysis changes or the size of one of the dimensions changes.

The GrassGro Analysis is an encapsulation in the software of the traditional process which for many DSS tools is a manual one. Building and executing a factorial experiment and then combining the results from all the treatments into a report centred on a chosen management question is done in a single step within GrassGro. The software manages the initial values for each scenario, how they are executed, where the results are stored, and how the results are displayed.

The ability of GrassGro to examine management questions through scenario analysis moves it beyond being just a system simulator (McCown, 2002) and into being a more complete DSS tool.

4.1. Reporting System

To obtain the most benefit from a well engineered DSS a good reporting system is essential. Many systems support examination of discrete output variables and also offer the ability to export them to third party tools. This is not an efficient process and can easily lead to confusion of the intention for the initial consultation. Tools such as HydroLOGIC (Richards, et al., 2008) and GrazFeed demonstrate compound reports. HydroLOGIC includes the presentation of multiple scenarios in a single report through the use of tabular data. To support the factorial analyses in GrassGro, the new reporting system needed to easily adapt to different types of input sources, use more than one source for a single report, present multiple report sections such as tables and charts in the same report, adapt to data sets from a multi-dimensional analysis, aggregate and transform data, and be easy enough for an experienced user to edit a report template.

The data aggregation and transformation abilities built into the GrassGro reporting system include probability distributions, percentiles, standard deviation and long term averages. Report sections can be constructed with fine-grained control over the aggregations applied to each data series within a transformation.

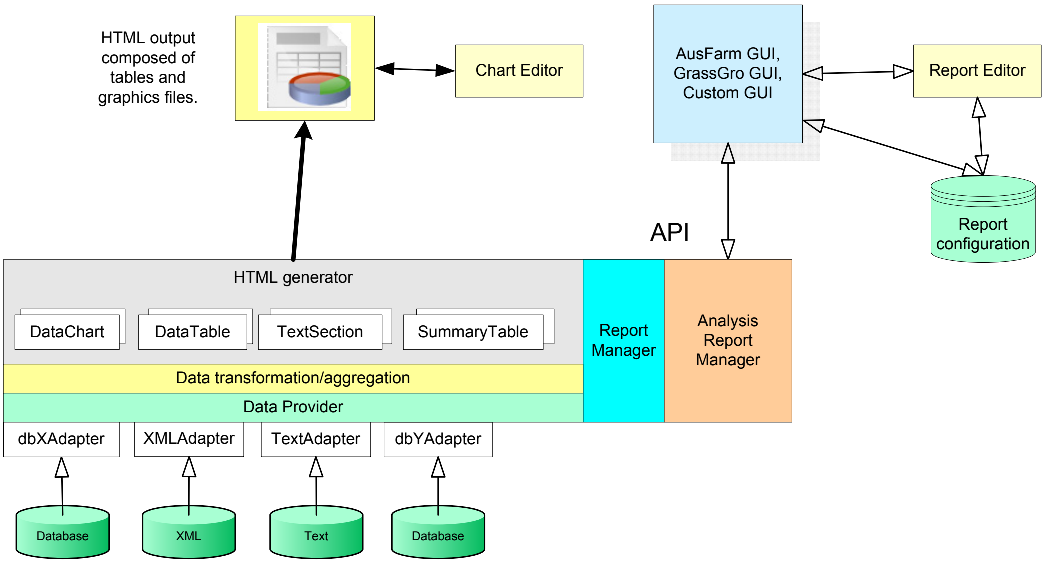


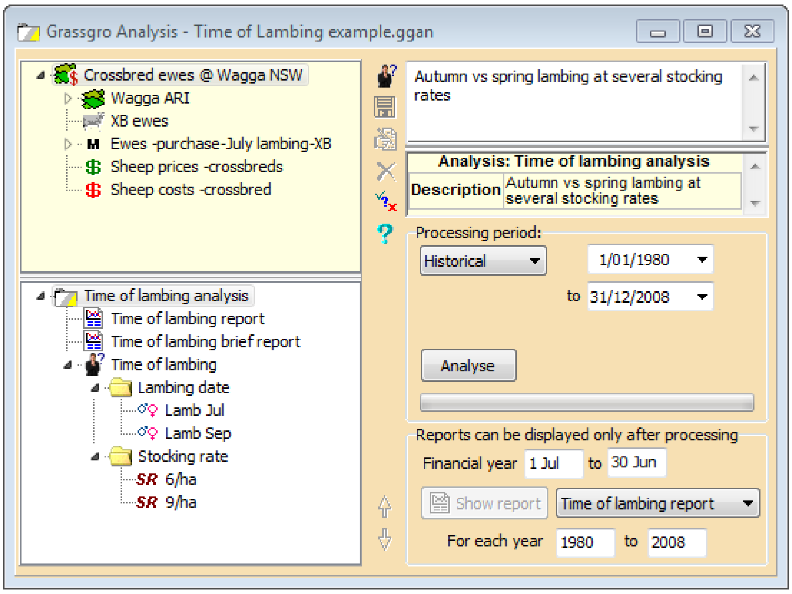
Figure 3. Logical view of the reporting system used in GrassGro.

The main purpose of the reporting system is to apply a report design or template to a number of result sets. The report template is able to be customised before being used by the Analysis Report Manager. The Analysis Report Manager, as shown in Figure 3, takes the structure of the factorial experiment and uses this to construct the appropriate layout of the final report. The report is made up of different report sections such as the DataChart and DataTable. Each of these sections draws its data from a Data Provider instance that can be specialised to extract data from a data source with a specific format.

The report is generated as a HTML document that displays standard HTML elements. The HTML report is not purely a static document. When the report contains charts the user can click on any of them with the mouse and open an editor that allows reformatting of the chart.

4.2. Factorial Analyses

To illustrate the way that a multi-dimensional analysis is constructed in GrassGro, Figure 4 shows an Analysis window with a two-dimensional experiment.


Grassgro Analysis - Time of Lambing example.ggan

- Crossbred ewes @ Wagga NSW
  - Wagga ARI
  - XB ewes
  - Ewes -purchase-July lambing-XB
  - Sheep prices -crossbreds
  - Sheep costs -crossbred

- Time of lambing analysis
  - Time of lambing report
  - Time of lambing brief report
  - Time of lambing
    - Lambing date
      - Lamb Jul
      - Lamb Sep
    - Stocking rate
      - SR 6/ha
      - SR 9/ha

Autumn vs spring lambing at several stocking rates

Analysis: Time of lambing analysis
Description: Autumn vs spring lambing at several stocking rates

Processing period:
Historical 1/01/1980 to 31/12/2008

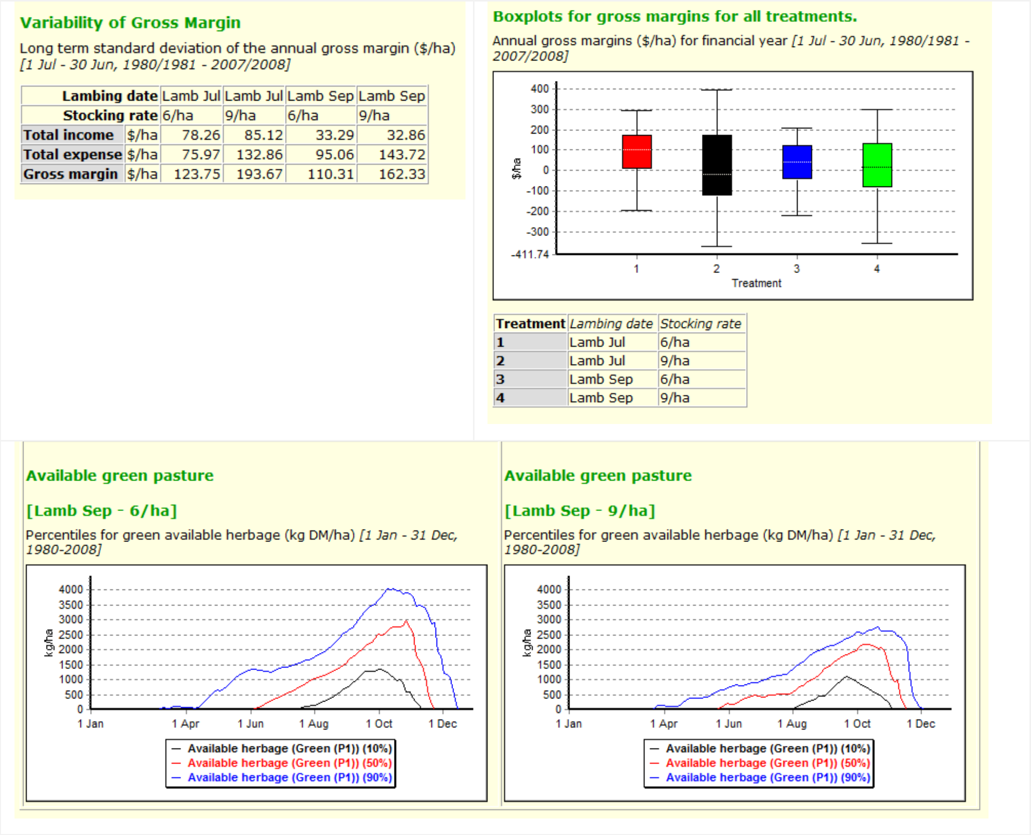
Analyse

Reports can be displayed only after processing
Financial year 1 Jul to 30 Jun
Show report | Time of lambing report
For each year 1980 to 2008

Figure 4. Analysis Window

The top left tree in the window displays the base Farm System and the Analysis is configured in the bottom left tree. This two-dimensional Analysis has Lambing date and Stocking rate Factors, each with two ‘values’. The Lambing date values are Reproduction components that consist of a number of related quantities. When this Analysis is processed, the base Farm system will have the matching components substituted with Factor values shown in the Analysis tree. This two dimensional Analysis will generate four sets of output results for use by the reporting system. Modifications to the factorial design are done using drag and drop mouse operations on the form and specialised component dialogs are used for editing of the factor values.

The report that is generated from this Analysis includes a number of charts, tables, summary tables and text sections (Figure 5). The combination of report sections in the report gives a user a comprehensive picture of the simulation experiment. Differences between scenarios can be identified easily. Economic risks are compared in the first part of the report and supported by the biophysical detail shown further down. After examining the report it is a simple process to return to the Analysis window and modify any of the proposed scenarios.



Variability of Gross Margin

Long term standard deviation of the annual gross margin ($/ha) *[1 Jul - 30 Jun, 1980/1981 - 2007/2008]*

Lambing date		Lamb Jul	Lamb Jul	Lamb Sep	Lamb Sep
Stocking rate		6/ha	9/ha	6/ha	9/ha
Total income	$/ha	78.26	85.12	33.29	32.86
Total expense	$/ha	75.97	132.86	95.06	143.72
Gross margin	$/ha	123.75	193.67	110.31	162.33

Boxplots for gross margins for all treatments.

Annual gross margins ($/ha) for financial year *[1 Jul - 30 Jun, 1980/1981 - 2007/2008]*

Treatment	*Lambing date*	*Stocking rate*
1	Lamb Jul	6/ha
2	Lamb Jul	9/ha
3	Lamb Sep	6/ha
4	Lamb Sep	9/ha

Available green pasture

[Lamb Sep - 6/ha]

Percentiles for green available herbage (kg DM/ha) *[1 Jan - 31 Dec, 1980-2008]*

Available green pasture

[Lamb Sep - 9/ha]

Percentiles for green available herbage (kg DM/ha) *[1 Jan - 31 Dec, 1980-2008]*

4000
3500
3000
2500
2000
1500
1000
500
0
kg/ha
1 Jan
1 Apr
1 Jun
1 Aug
1 Oct
1 Dec
Available herbage (Green (P1)) (10%)
Available herbage (Green (P1)) (50%)
Available herbage (Green (P1)) (90%)

Figure 5. Some of the typical sections that are displayed in a report

The System Designer, who designs the Issue, edits the report templates through a dialog that displays each report section in a hierarchical tree. Report sections can be enterprise specific and can be repeated in the layout for every treatment in the Analysis. A report section such as a single chart or summary table can also display all treatments. When building reports, sections can be copied and pasted between reports and a library of pre-prepared template sections is also available.

5. DISCUSSION AND CONCLUSION

DSS systems can provide valuable insight into the behaviour of systems. However without the ability to easily examine a range of alternative scenarios, the user can remain poorly informed when solving their management question. The scenario analysis process developed for GrassGro has addressed many of the concerns when using a DSS to investigate a management question for a complex system. Much of the redundancy, time costs and potential for error have been reduced by building a software system that manages the analysis process.

To build a DSS that supports factorial analyses, the system model must be decomposed effectively into sub components. Factorial analyses in GrassGro rely on the Farm System being decomposed into components that are sufficiently decoupled so that dependencies between them are kept to a minimum. This is one of the challenges in building a software application that can support the demonstrated type of factorial experiment.

Constructing efficient and meaningful Analyses in GrassGro also requires a good understanding of the factor components and how they can best be used to address a management question.

An integrated reporting system to summarise and display the results of an Analysis is another crucial element. The reporting system in GrassGro version 3 has greatly improved the ability of a DSS to address the management questions posed by graziers and producers.

GrassGro has been demonstrated to be an effective tool for agricultural advisors in departments of agriculture. For example, the ability for GrassGro to quickly examine a set of scenarios has meant that it is a useful tool for climate change scenario analysis. The current DAFF, MLA, Dairy Australia and AWI project *Climate Change Adaptation in Southern Australian Livestock Industries*, is using GrassGro to help farmers think about the future by illustrating numbers of possible climate scenarios.

The design of the GrassGro Analysis implementation has been introduced into the more advanced generic modelling tool AusFarm. Some of the concepts have also been recently introduced into APSIM that allow it to be used for building factorial runs of simulations. The approach taken in the GrassGro design is robust and useful for other systems and benefits can be obtained by building a system with these features.

REFERENCES

Cook, F.J., Jordan, P.W., Waters, D.K., Rahman, J.M., (2009). WaterCAST – Whole of Catchment Hydrology Model An Overview. 18th World IMACS/MODSIM Congress, Cairns, Australia 13-17 July 2009.

Donnelly, J.R., Freer, M., Moore, A.D., (1998). Using the GrassGro Decision Support Tool to Evaluate Some Objective Criteria for the Definition of Exceptional Drought. *Agricultural Systems* 57, No.3, 301 – 313.

Donnelly, J.R., Freer, M., Salmon, E.M., Moore, A.D., Simpson, R.J., Dove, H., Bolger, T.P., (2002). Evolution of the GRAZPLAN decision support tools and adoption by the grazing industry in temperate Australia. *Agricultural Systems*. 74, 115–139.

Freer, M., Moore, A.D., Donnelly, J.R., (1997). GRAZPLAN: decision support systems for Australian grazing enterprises. II. The animal biology model for feed intake, production and reproduction and the GrazFeed DSS. *Agricultural Systems* 54, 77–126.

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., Smith, C.J., (2003). The Agricultural Production Systems Simulator (APSIM): its history and current capability. *European Journal of Agronomy*. 18, 267 – 288.

Mahmoud, M., Liu, Y., Hartmann, H., Stewart, S., Wagener, T., Semmens, D., Stewart, R., Gupta, H., Dominguez, D., Dominguez, F., Hulse, D., Letcher, R., Rashleigh, B., Smith, C., Streetm, R., Ticehurst, J., Twery, M., van Delden, H., Waldick, R., White, D., Winter, L., (2009). A formal framework for scenario development in support of environmental decision-making. *Environmental Modelling & Software*. 24, 798–808.

Moore, A.D., Donnelly, J.R., Freer, M., (1997). GRAZPLAN: decision support systems for Australian enterprises. III. Pasture growth and soil moisture submodels, and the GrassGro DSS. *Agricultural Systems*. 55, 535–582.

Moore, A.D., Holzworth, D.P., Herrmann, N.I., Huth, N.I., Robertson, M.J., (2007). The Common Modelling Protocol: A hierarchical framework for simulation of agricultural and environmental systems. *Agricultural Systems* 95, 37–48.

Perillat, B.J., Brown, W.J., Cohen R.D.H., (2004). A risk efficiency analysis of backgrounding and finishing steers on pasture in Saskatchewan, Canada. *Agricultural Systems* 80, 213–233.

Richards, Q.D., Bange, M.P., Johnston, S.B., (2008). HydroLOGIC: An irrigation management system for Australian cotton. *Agricultural Systems*. 98, 40–49