The development of a farming systems model (APSIM) a disciplined approach.

Dean Holzworth

Holger Meinke

Peter DeVoil

Malcolm Wegener

Neil Huth

Presenter/Author Information
Dean Holzworth, Holger Meinke, Peter DeVoil, Malcolm Wegener, Neil Huth, Graeme Hammer, Mark Howden, Michael Robertson, Peter Carberry, David Freebairn, and Chris Murphy

The development of a farming systems model (APSIM) – a disciplined approach

Dean Holzworth [1], Holger Meinke [2], Peter DeVoil [2], Malcolm Wegener [3], Neil Huth [1], Graeme Hammer [4], Mark Howden [5], Michael Robertson [7], Peter Carberry [1], David Freebairn [6] and Chris Murphy [2]

[1] CSIRO Sustainable Ecosystems / APSRU, P.O. Box 102 Toowoomba, Qld, 4350.
[2] Department of Primary Industries and Fisheries / APSRU, P.O. Box 102 Toowoomba, Qld, 4350.
[3] Agricultural Production Systems Research Unit, School of Natural and Rural Systems Management, The University of Queensland Brisbane, Qld, 4072, Australia
[4] Agricultural Production Systems Research Unit, School of Land and Food Sciences, The University of Queensland Brisbane, Qld, 4072, Australia
[5] CSIRO Sustainable Ecosystems, GPO Box 284, Canberra, ACT, 2601.
[6] Department of Natural Resources, Mines and Water, NRMW, APSRU, P.O. Box 318, Toowoomba, Qld, 4350.
[7] CSIRO Sustainable Ecosystems / APSRU, Private Bag 5, PO Wembley, WA 6913.

Dean.Holzworth@csiro.au

Abstract: The Agricultural Production Systems Simulator (APSIM) is a mature and stable modelling framework used widely in Australia and elsewhere in the domain of farming systems research and extension. It is capable of simulating a diverse range of farming systems including broadacre dryland and irrigated cropping, small holder farming and on-farm agroforestry systems. This includes the interaction of trees and crops and, through collaboration with other groups, integrated stock and cropping enterprises. APSIM was developed primarily as a research tool to investigate on-farm management practices especially where outcomes are affected by variable climatic conditions. Its use has been extended to looking at modifying farm practices and to include analysis of natural resource management issues including salinity and solute movement, climate risk studies, and climate change scenarios to name but a few. In recent times commercialisation activities have increased to broaden the user base by taking the utility of APSIM directly to consultants and farmers and tailoring it to their needs. This paper details APSIM's evolution over the past 15 years including its conception, specification, construction, performance and use. APSIM's development was very much a collaborative effort between multiple organisations and it has been driven by both user needs and scientific advances in crop and soil science, while the implementation of these advances in modelling was carried out by professional programmers and software engineers to assure best practice and quality control. APSIM's development is on-going. Here, we chronicle the development effort to date and detail some of the lessons learnt along the way.

Keywords: Farming systems modelling; APSIM; simulation model

1. INTRODUCTION

In Australia the need for interdisciplinary, agricultural systems R&D became obvious during the 1980s. From those early attempts to model agricultural cropping systems, the Agricultural Production Systems Research Unit (APSRU) was created in 1991 as an inter-agency, interdisciplinary research group with the explicit aims to a) facilite research collaboration, b) co-develop and manage research tools, methods and resources, and c) influence agricultural systems research and design processes. The formation of APSRU provided the stimulus for the development of APSIM, addressing the need for modelling tools to provide reliable estimates of crop production and assessments of associated environmental consequences in relation to climate, soil, management and genetic characteristics of a range of crops [Keating *et al.*, 2003]. This paper outlines the bringing together of several disciplines to create a robust, flexible simulator for agricultural systems.

Jakeman *et al*. [2006] defines 10 important steps for disciplined model development. While we acknowledge that these steps are useful in helping to shape model development and have, indeed, strongly influenced the implementation of APSIM, its design was and is guided by other principles. APSRU's emphasis has always been on the effective description of processes (both biophysical and management) rather than on extensive parameter value identification and estimation techniques.

Hence, this paper is structured around the concepts of Conception/Specification, Construction, Model performance and testing, and Model Delivery, each with sub topics. Note though, the first four sub topics are the same as the first four proposed by Jakeman et. al.

2. CONCEPTION AND SPECIFICATION

2.1 Purpose for building APSIM

At the outset, three key issues drove the development of a dynamic systems simulation capability to address the short and long-term consequences of crop management; quantifying the dynamics of genotype x management x environment interactions (GxMxE); and the use of modelling as a communication tool between disciplinary groups. Models have also been useful as a depository of accumulated knowledge about processes operating in agricultural ecosystems (eg. plant growth, soil physics etc). All three issues are still valid today and they are the key driving forces behind the on-going development of APSIM:

1. **Short versus long-term consequences of crop/cropping systems management**. When APSRU was formed in 1991, decision makers as well as scientists were asking questions about the overall performance of farming systems (eg. optimal crop rotations, fallow management and its impact on stored water reserves, impact of residues on nitrogen and soil organic matter cycles and subsequent growth of crops [Muchow and Bellamy, 1991]. To answer such questions required the ability to correctly simulate intra-seasonal responses to variability in environmental conditions (eg. seasonal conditions, nutritional status) as well as longer term soil processes such as organic matter decline, soil erosion, structural degradation, soil acidification. None of these questions could be addressed with existing modelling tools that either considered single-crop issues e.g. the CERES family of models [Jones and Kiniry, 1986], Wheatman [Woodruff, 1992], QSUN [Chapman *et al*., 1993], or the IBSNAT project [Uehara and Tsuji, 1991]. In addition they lacked the detailed, dynamic representation of crop/soil interaction needed for the evaluation of long-term management options eg. CENTURY, [Parton *et al*. 1987], EPIC [Williams 1983], or PERFECT [Littleboy *et al*. 1989]. Although strongly grounded in many of the same concepts, APSIM was designed originally as a farming systems model (albeit with a near exclusive focus on crops and crop rotations in the early years), rather than as a crop model, thereby allowing assessments of short- as well as long-term systems dynamics. Specifically, the initial development of APSIM focused on quantifying the carry-over effects from one crop/fallow period to the next.

2. **Quantifying the dynamics of genotype x management x environment interactions (G*M*E)**
Initially issues associated with risky crop management decisions – crop and cultivar choice, planting time, density and fertility options - were the primary targets of crop modelling activities using APSIM. They remain a key feature of APSRU activities and have now reached a level of maturity where commercial partnerships have been formed to progress delivery (see below). In addition, recent advances in biotechnology raised questions about how we could improve crop design to achieve both productivity and sustainability outcomes. Plant breeders and physiologists wanted to enhance the efficiency and effectiveness of plant breeding and crop improvement programs against the background of extremely high climatic variability in Australia and a wide range of proposed management options [Shorter *et al*., 1991], [Hammer *et al*., 2002]. Novel gene-to-phenotype crop simulation platforms were needed to meet this objective. This required a crop-modelling framework tailored to the needs of gene-to-phenotype simulation in order to quantify the physiological and genetic basis of genetic variation in key crop traits and quantify the G*M*E influences on phenotypic variation. This created tensions between 1) the reliability of simple, descriptive approaches that necessarily have limited capacity for extrapolation beyond the data used in their construction and 2) the need for a more detailed, mechanistic representation of the physiological processes involved in crop growth able to be extrapolated to new situations but where the predictive ability of the model can be reduced due to a (potentially) large increase in the number of parameters. For effective use in exploring G*M*E interactions it was clear that process based phenomenological

descriptions of the dynamic processes involved in the complex activity of plant growth and development were needed and consequently that the APSIM modelling framework must be able to accommodate various levels of complexity, depending on the intended application. The development of the generic crop template (discussed below) arose partly in response to this need.

3. **Modelling as a communication tool and depository of knowledge across disciplines**. Traditionally scientists have used simulation models as 'knowledge depositories' in order to describe an area of interest [Meinke *et. al.* 2005]. Systems models such as APSIM take this concept further by providing a means for effective communication across all the disciplines involved to address issues affecting the target systems – in our case, farming and agriculture. The dual function of a systems model as a depository of knowledge and as a tool to facilitate communication of scientific knowledge across disciplines also makes it an ideal teaching tool at graduate and particularly post-graduate level.

2.2 Specification of the context: scope and resources

Given the key objectives outlined in the previous section, APSIM needed to simulate the main biological and physical processes in a farming system in sufficient detail to allow meaningful analysis of the problem domain. Clearly, it required the following features:

- A long-term systems focus with carry-over effects from one crop/fallow period to the next.
- Flexible management capabilities to facilitate the construction of a broad range of simulations that closely resemble management actions taken by farmers.
- A framework capable of capturing advances in science and facilitating communication between the various disciplines.
- To re-use, where possible, prior knowledge and existing models.

Figure 1 illustrates the farming system that APSIM was initially designed to simulate. The soil is the central focus of the system: crops, weather, and on-farm management come and go, finding the soil in one state and leaving it in another. There is an equal emphasis, and level of detail, on above ground and below ground processes.

The identified key processes were:

- growth and yield of crops, pastures, and more recently, trees
- soil processes including evaporation, runoff, water, nitrogen and phosphorus movement, and changes in pH
- surface residue dynamics and erosion
- a diverse range of crop and farm management options

In 1991, when APSIM's development commenced, two software developers and six scientists were actively engaged in model design and implementation. This has since grown into a software engineering group (SEG) with six software developers and ten scientists contributing to the ongoing development of APSIM. In

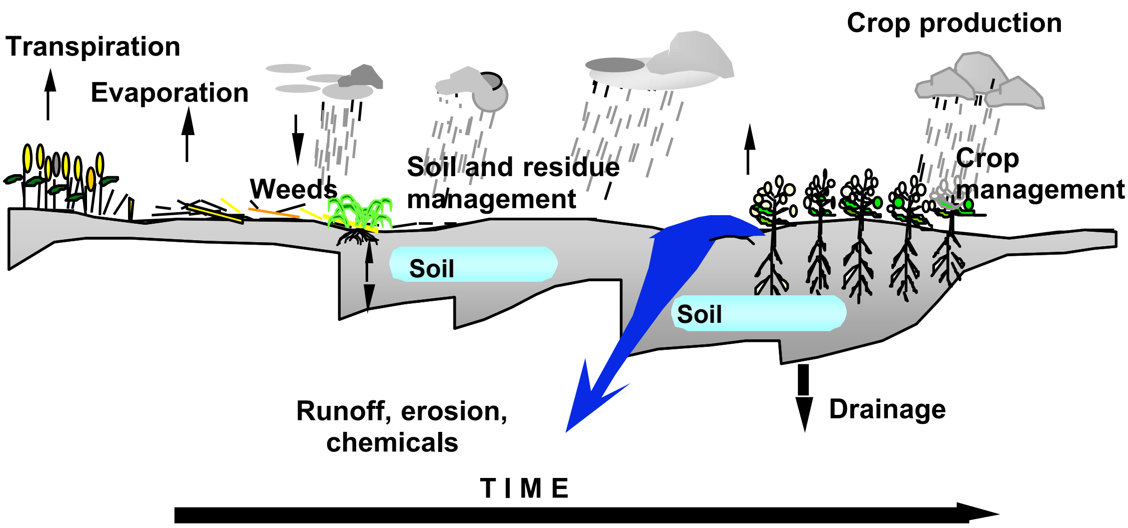

Figure 1: A diagrammatic view of the problem domain addressed by APSIM

addition, other scientists from many institutions around the world continuously contribute to the advancement of the science within APSIM. However, the responsibility for the accurate implementation of this science within the APSIM framework rests with the SEG.

2.3 Conceptualisation of the system, specification of data and other prior knowledge

Initial conceptualisation of the farming systems model used concepts and components from existing soil and plant system models where they were available. The first APSIM crop modules, for instance, were based on CERES-Maize [Jones and Kiniry, 1986], and existing soybean [Sinclair, 1986] and wheat models [van Keulen and Seligman, 1987], which had already been tested and enhanced by APSRU scientists [Carberry *et al.*, 1989]; [Hammer and Muchow, 1991]. Similarly, for aspects of the soil water balance and systems effects on soil erosion, existing concepts and components of other models [Ritchie, 1972] as well as our own innovations [Littleboy *et al.*, 1992] were used.

From this starting point, improvements were made in all modules as knowledge and process understanding accumulated - APSIM became a 'knowledge depository'. Routines and parameter values were enhanced on the basis of experimental evidence from studies of our own and others. In this way, the soil water and nitrogen routines were progressively developed by Probert *et al.*, [1998c] and the crop routines were generalised leading to the development of a generic crop module template as described by Robertson *et al.* [2002] and Wang *et al.* [2003].

Given that these early models provided the level of detail and understanding needed to meet the demands identified earlier in this paper, it seemed logical to have APSIM operate on a similar level of detail. Like the previously listed models, APSIM also operates on a point (field) scale using a daily time-step. However, unlike other models, APSIM required a management interface, i.e. a flexible approach to reproducing complex on-farm management practices: fertiliser and irrigation scheduling, crop sequencing, fallow management and intercropping to name just a few. Real-world, on-farm management practices are complex and dynamic and APSIM needed to reflect this.

Wrapped around this core set of user requirements was the need to implement APSIM with longevity in mind. Rather than designing a system that only addressed immediate needs, we needed a concept that could guarantee 'business continuity' in the long term. Hence, the decision was made to strategically invest in a modular, extendable framework that could grow over time and be tailored to future needs. Such a strategic approach is rare, given the tight and short-term funding environment that applied when APSIM was created and still applies now.

A decision was also made to implement an industry-standard software engineering process in order to ensure quality control of science implementation. This needed to happen in a science-focussed way and while a painful process to begin with, led to the development of a sound relationship between the software developers and scientists.

3. CONSTRUCTION

3.1 Selection of model structure

The above user requirements led to the modular, component-based modelling framework (Figure 2).

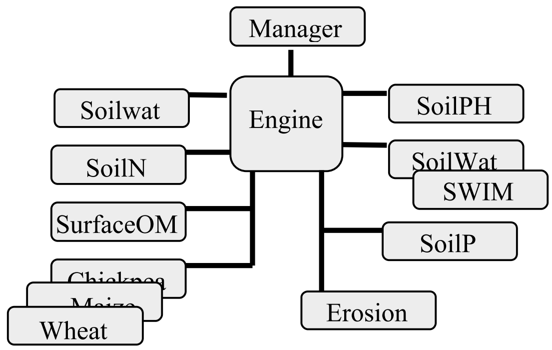

Figure 2: A conceptual diagram showing the relationship between APSIM's modules and the engine.

This early design involved a conceptually simple, centralised engine into which *modules* could be connected. Each module provides a small piece of simulation functionality with the 'engine' coordinating the flow of data/variables between the modules. This design allowed an incremental approach to development. The engine was developed first and then modules slowly developed over time, building our capability to the point where the current version of APSIM (V5.0) contains approximately eighty modules with new ones being continually developed. Many of these modules have been published as journal papers over the years, for example; APSIM's generic crop template described by Wang *et al.* [2003];

the SoilWat module published by Probert *et al.* [1998c] and specific crop routines published by Hammer and Muchow [1991] and Robertson *et al.*, [2002] .

In order to create a modelling framework that was both useable and useful, as defined by Huth *et. al.* [2005], modules have been designed to capture the major elements of the problem domain in a manner that can be quickly and efficiently described by a model user e.g. Keating *et al* [2002]. This enables the modeller to specify a system using a wide array of problem domain objects including crop and soil components. Inherent in the modular approach is the provision of appropriate parameters. An important design criterion was to use (to the maximum extent practicable) parameters where the necessary values can be easily obtained, either from the literature, or experimentally. This has led to the establishment of comprehensive databases of crop physiological measurements (REMS), described by McLean *et al.,* [2004] and soil physical characterisations by Dalgliesh *et al.,* [1998] for most of Australia's important crops and cropping regions. In addition, there have been a significant number of cases where model performance has been checked and verified for overseas countries.

Where data exists for the major areas of model input, the model is designed to make these data sources immediately available. For example, long term daily climate information is available for much of the Australian continent [QCCA, 1998]. APSIM modules have been explicitly designed to make use of this readily available daily data rather than relying on detailed diurnal data that cannot be easily obtained.

One module of particular importance is the *Manager* module. It allows the modeller to provide script, in the form of a programming language, which exactly specifies the timing and detail of farming operations. By externalising farm management decisions with a very flexible scripting language, modellers have a large degree of flexibility to develop a wide variety of simulations ranging from simple field experiments to mimicking the decision-making processes of a real-world farm manager.

The decoupled, decentralised design also provided a mechanism for scientists to implement their 'knowledge' independently but within a consistent framework. They can develop new modules or work on existing modules without risk of interfering with each other and, indeed without the need to fully understand the science contained in some of the other modules. This modular approach also allows direct comparisons of specific processes (e.g. soil water dynamics) without the confounding factors of different implementation of other processes (such as crop growth) that occur with model-to-model comparisons. Modules still need to communicate as all modules require data from elsewhere in the system, but at least the software issues are solved by this design. The approach that worked well for many years was for scientists with programming skills to develop the modules, but more recently, due mainly to time pressures and quality of source code, software developers have taken over the role of implementing the science within APSIM. This is now done by scientist/programmer pairs jointly working on new science functionality.

As testament to the flexibility of the original design, the APSIM model was recently revised to provide a multi-point (multiple field) capability by simply allowing multiple instances of the engine and modules for multiple points to run simultaneously. An instance of the engine and some modules are also created at the farm level to provide over-arching management of all fields, thus introducing hierarchy into the design (Figure 3).

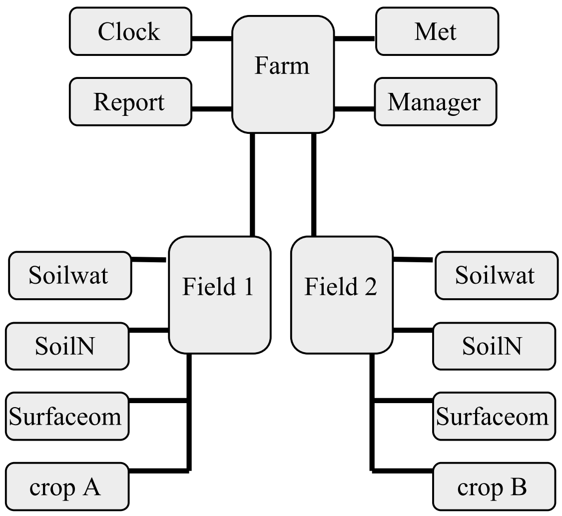

Figure 3: The current conceptual diagram of APSIM.

As part of this work, APSRU together with CSIRO-PI co-developed a 'Common Modelling Protocol' (CMP) that defines the way data flows from one module to another [Moore, 2001]. This software protocol allows module exchange between different compatible modelling environments. For APSRU, it provides access to the CSIRO Plant Industry FarmWi$e model [Donnelly *et al.*, 2002], a suite of pasture and livestock components. This capability has already been used in research projects, allowing sheep and cattle to graze crops and pastures simulated by APSIM. From a software developer's perspective, the pleasing aspect was that the science in APSIM's modules did not need to change upon implementing the protocol: only the infrastructure code required changes.

3.2 A bit of software process aids in model development

It was clear from the outset that a new emphasis on sound software engineering principles was needed as multiple software developers and scientists all contributing to the source code of APSIM could create software chaos.

The development process for APSIM could best be described as 'agile' rather than 'traditional'. For several years, the *Agile* software movement has been promoting simple, client focussed processes for constructing software applications (see http://agilemanifesto.org/). In fact many of the software processes used by the SEG originated from the Agile software movement. As well as the more traditional use of version control, the emphasis is on automated nightly testing (see section 4.2), small frequent releases of modified software and iterative/evolutionary development. Working software is given priority over developing comprehensive documentation and developing data standards. Software is designed to respond to changing requirements rather than developing extensive case scenarios in the hope of capturing all possible requirements. The SEG is still involved in many of these lower priority processes, but the emphasis is on the source code.

To summarise the SEG philosophy:

- APSIM is always kept in an operational state. By doing this, the SEG can release updated versions of APSIM any time.
- New modules based on evolving science are always developed incrementally. Given the amount of time it takes to gain user's confidence, the SEG are very reluctant to discard existing operational modules and start afresh. It is always much better to rework existing functionality to incorporate new ideas and new science than to build something completely different.
- Tied up with the small iterative style of development, the SEG always commits source code changes to version control as frequently as possible, never having code 'locked' for long periods of time.
- The nightly automated testing regime, as outlined in section 4.2, is a critically important feature of the quality control process.

3.3 More than good software: capturing the science

The third issue outlined in section 2.1 discusses the requirement for a dynamic framework for capturing scientific advances in process understanding while retaining and improving predictive capacity. APSIM's modular approach to software design and the software engineering principles outlined in the previous section facilitate this capacity. However, the development process also requires co-operation among scientists, software engineers, and model users. Co-ordinating procedures are required to manage sometimes conflicting requirements for improved science content, software management, and stability for users.

The approach taken to the on-going development of the object-oriented generic plant module (oo-Plant) in APSIM is an example of such co-ordination. Interested scientists, software specialists, and model users meet regularly to progress design and implementation issues. This provides a forum for quality assurance of any new process science about to be incorporated into the models. Software design to maintain a stable platform can be considered while accommodating the range of users' needs. There can be discussion about any concerns relating to predictive capacity of an application, as well as prioritisation of effort and management of human resources.

Given the flexibility of the APSIM platform, and the range of skills and ideas brought to this science-software-application interface, it is usually possible to find a path acceptable to all. In some instances this may involve representing some processes at various levels of detail and allowing optional use of the resultant sub-modules for different applications. For example, for some plant breeding applications, the conventional coefficient-based photo-thermal phenology algorithm can be replaced seamlessly by an optional routine connecting back to a representation of the controlling gene network [van Oosterom *et al*, 2004].

3.4 Model calibration: How are parameter values found?

Large detailed systems models such as APSIM place a heavy requirement on the level of input data required. Without the requirement of building modules that use real, measurable parameter values, parameterisation and calibration would be extremely problematic. Model users are encouraged to use published methods for parameter measurement or derivation in order to greatly reduce the calibration effort. Databases of previously tested crop and soil specifications are also provided with the model distribution.

Model developers, however, often require formal methods to optimise parameter sets for distribution to users. In these cases, optimisation tools can be used in conjunction with extensive experimental datasets. However, this is only ever applied in a manner where controlled experiments allow for the study of a single component of a model where other factors have been carefully removed in the experimental design. A common example is the modelling of phenological development in plants e.g. Robertson *et al.* [2002]. Many of these model components are not dependent upon site or season and so are provided to users, thus further reducing the level of calibration required of the user.

Our experience has indicated that the provision of carefully tested parameter sets and methods for determining site-specific values can significantly reduce the calibration requirement of the model user. Whilst this may not be so in other problem domains, it appears to be very successful in the area of agriculture or natural resource management.

A note of caution is required though. We should not think that the result will be either sensible or stable in situations far beyond the range of these parameter sets. In fact, the wide array of possible parameter set values will most likely result in outputs that will be difficult to reproduce by different operators. This is not a criticism of models like APSIM, rather a statement of reality. If a user does not have a feel for sensibility in any system being simulated, then it is unlikely that any results will be in the sensible range either.

4. MODEL PERFORMANCE AND TESTING

4.1 Model validation: How does the model perform?

New users of APSIM (scientists, software engineers, agricultural consultants, and farmers) invariably ask questions such as “how reliable is the model”? On deeper examination, this usually turns out to be questions not directly related to APSIM, but about models in general. However, sensitivity surrounding model validation usually has its roots in who is asking the question. Typically discord arises when one person’s

Source of uncertainty	Steps to deal with this uncertainty
Processes not included in the modelling framework (e.g. impacts of pest and disease impacts)	Incorporation of appropriate modules (e.g. a rodent damage module; Brown *et al.,* 2006)
Processes not included in specific simulation runs (e.g. not including weeds in a crop fallow)	Diagnostics that identify the likely source of error and suggest ways to establish a better description of the system of interest
Errors in parameterisation	Continuing development of parameter sets.
The structure of the model (e.g. where a daily time step, canopy model is inappropriate)	Consider using a different timestep (using the Clock module), assess the error from the underlying assumptions (e.g. where there is a sparse canopy). In some cases incorporation of other alternative approaches
Aggregation of diverse processes or factors (e.g. lumping soil parameters which may be non-linear in their effects)	Represent the diversity where possible (e.g. use the multi-point capability to represent different soil types in a paddock) otherwise recommend alternative approaches if error is unacceptable.
Simplification of complex processes (e.g. phenology, photosynthesis or transpiration)	Continuing re-evaluation of alternative representations and provision of options to use these (e.g. use of either the conventional phenology algorithm or one based on the controlling gene network; section 3.3)
Errors arising from stochastic variation in the environment interacting with the management regime and biophysical processes (e.g. different results from rotations depending on the start year)	Routines that vary the start date of simulation runs so that all possible combinations of start date are used and taking a composite value of these when comparing farming systems.
Lack of software integrity	Effective quality control (see 4.2)

perspective of the models usefulness disagrees with another.

Models are inherently an incomplete description of the systems they are intended to represent. Consequently there is always a degree of uncertainty in the results that they provide. This uncertainty can be derived from many different sources (see Table 1 for examples). APSRU is committed to incrementally reducing those sources of uncertainty where it is practicable and where it is required by the user community. The approach is to explicitly acknowledge the maxim 'All models are wrong, some models are useful'.

A common method of estimating the uncertainty of either the entire modelling framework or of specific science modules is to undertake a process termed validation: where model output of the system component of interest is compared with measurements of that component. The theoretical basis of validation is outlined well by Oreskes [1994], who states that model validation does not establish 'truth'; rather it shows that the model behaves consistently with some set of observations. It follows that each user group will have different standards (observations) by which to measure model performance. 'Science' users are comfortable with some combination of predicted:observed plots (e.g. Figure 4), and measures of error. However, in some cases this requires augmenting with consideration of degree of biophysical rigour in process representation.

'Ordinary' users may have other criteria related to uncertainty pertinent to their application of the model. Huth *et al.* [2005] describe these as Sensibility Tests – tests that demonstrate the model in general behaves in accordance with their understanding of the problem domain.

Software Engineers have yet another perspective on valid operation: Does it crash? Does it pass unit tests? Does it do what it did yesterday?

Each 'biological' module within APSIM has passed a scientific level of acceptance; usually in the form of a published, peer reviewed paper or technical document. For the science users of the model, these papers answer most of the questions of validity, as they describe both the module's internal operation (how it fits with our understanding of physical processes) and how the model behaves compared to experimental data – a procedure outlined by Rykiel [1995]. Operationally, the software maintenance process described in section 4.2 ensures the model does not deviate from this published material.

In the same manner, 'infrastructure' components of the model (eg engine, reporting, management) pass a suite of unit tests and also support the science tests outlined above.

Recently, the SEG have moved towards automated production of 'validation plots' where graphs of predicted:observed data and associated statistics of error are produced as part of the daily testing process, and published on APSIM's web site (www.apsim.info) alongside supporting documentation and source code. However the sheer volume of these images (~600 images just concerning wheat) is preventing meaningful interpretation and alternatives are being sought.

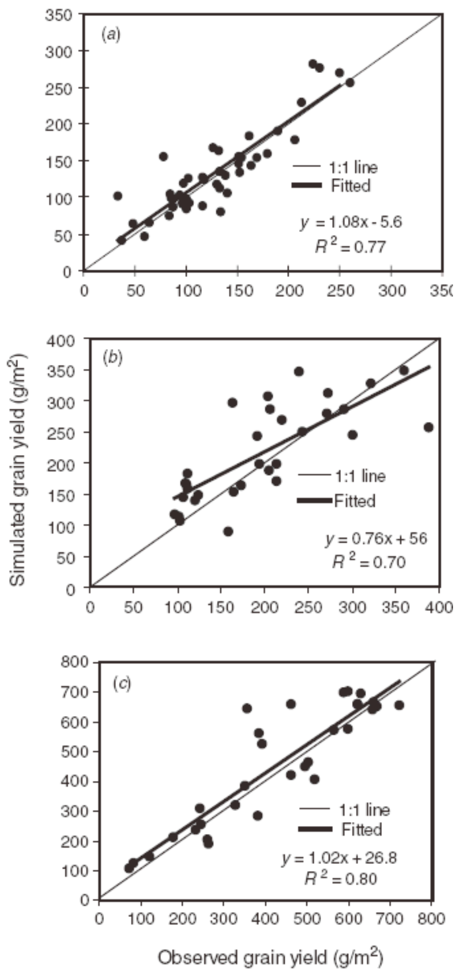

Figure 4: Observed and simulation grain yield for (a) mungbean, (b) peanut, and (c) chickpea [Robertson *et al.* 2002]

4.2 Ensuring that the model behaves reliably and sensibly

While scientists are primarily concerned with model validity and performance in terms of predicted:observed model accuracy, software developers are more concerned with reliability. To measure APSIM's reliability, an automated testing regime has been implemented that is run nightly or more frequently during peak periods of development. At each run of the test suite, all source code is extracted from version control, compiled and evaluated by executing a suite of simulation runs. These runs fall into several categories, each performing a different function:

- Unit tests operate at the lowest level and test individual functions/methods looking for expected behaviour.
- Test simulations reproduce defects that have been identified and fixed. They ensure that fixed defects don't reoccur.
- Reference sets exercise a module at the extremes of expected input data. For example, for crop modules, simulations of very low and high water and nitrogen content are created looking for unexpected behaviour.
- Traditional validation tests produce the familiar observed:predicted plots that are important for evaluating model performance.
- Sensibility tests simulate real-world conditions under a variety of environment and management scenarios but outside the range of the validation tests. Experts in the area of simulation (agronomics, consultants, growers) are then consulted to provide 'observed' outputs. The test results are then compared with these observed data thus providing a measure of model *sensibility*.
- User runs are simulations that users have supplied to exercise APSIM in unforeseen (by APSIM's developers) and different ways. By including these tests, the amount of source code covered by tests is increased and a degree of backwards compatibility with real-world user simulations is achieved. After all, users are the ultimate judge on how reliable and sensible a model is and so it is important that their 'real' simulations are included in the suite.

The underlying premise behind the entire testing framework is change. We specifically look for simulation outputs that change unexpectedly. Each morning, the results of the nightly testing are compared with known good outputs. Sometimes changed outputs are expected, for example if new science has been added, but at other times results are expected to be identical. In the former case, the new outputs may then become accepted as the new standard of 'good' outputs. When the unexpected happens though, it is always a priority to locate the 'misbehaving' source code and fix it. Priority is always given to a 'clean' build process before any further changes to source code can be committed to version control.

The model testing process described here was built in an evolutionary way and is ongoing. When a defect is found in APSIM, a reproducible test case is created, the defect fixed and then the test is added to the suite. When new science is added to a module, tests are created to exercise that new functionality. When a user creates a simulation that is different in some way, it is added to the suite.

The aim is to improve the chances of finding unexpected model behaviour – we do not claim to have an infallible testing process in place.

5. MODEL DELIVERY

5.1 How is the model used?

APSIM has been designed as a multi-purpose simulation platform. Originally, crop models were developed to deal with risky crop management decisions in the face of climatic variability. The models simulated plant growth and crop development in response to environmental inputs (water, temperature, solar radiation, nutrients) with the ultimate aim of estimating the yield of harvestable material from a commercial crop as precisely as possible. At the heart of these models is the relationship between crop yield and various inputs (climatic conditions such as rain, temperature and solar radiation, nutrients, and management interventions such as irrigation or fertilisation) that may or may not be affected by crop residues left on the soil surface from a previous crop. These residues can affect surface runoff, soil temperature, surface evaporation, and soil moisture and thus many processes that contribute to crop growth and yield as well as affecting the state of the environment in which the crop is being grown. This is where the need for good science arises so that the model simulates the processes appropriately and precisely, in ways that are easily computable, and the results are believable.

In addition to crop yield, models such as APSIM generate a large range of complementary output variables that can be very helpful in analysing resource management problems. Community concern about off-farm impacts of farm inputs such as nitrogen fertiliser has increased in recent years. Therefore, farm management practises that might cause long term resource degradation have come under close scrutiny. In Australia, problems such as increasing dryland salinity from large scale tree clearing are obvious in farming areas and there is much concern that pollutants from crop production such as pesticides, nitrogen in runoff, or nitrogen leaching into groundwater do not affect freshwater bodies or pristine marine environmental areas such as the Great Barrier Reef. These long term resource management problems cannot be investigated adequately with short term field experiments. Instead simulation models, like APSIM, can be used to quantify likely long term changes in soil and water resources or to evaluate the likelihood of pollution or degradation occurring. Modelled variables such as nitrogen leached below the root zone, deep drainage or changes in the level of the ground water table, can assist in evaluating the impact of various management practices on the state of farming resources.

However, simulation models that were essentially designed to simulate small areas of crop, while simulating environmental effects at the paddock scale reasonably well, need to be used carefully when evaluating broader-scale effects. There are both physical and economic limits to be considered when aggregating results from point-scale models and using them to evaluate landscape impacts.

APSIM was originally developed by researchers, for researchers, but over the past fifteen years it has also been widely used by consultants and indirectly, via derived products and model output presentations, by growers. It has been used:

1. **Scientifically**, to investigate and progress scientific issues (eg water use efficiency, trait evaluation etc).
2. **Analytically,** to investigate systems dynamics and interactions, generate new insight, test hypothesis and to quantify responses.
3. **Operationally,** to provide quantitative information relevant to stakeholders packaged as derived products eg Whopper Cropper [Nelson et al, 2002], Yield Prophet [Hochman et al, 2006].
4. **Practically,** as a tool to communicate with peers or stakeholders to stimulate and guide discussion.

How many users of systems models explore the inner workings of the model through exploration of individual processes. How well are models set up to do this with minimal effort? For example, it is usual for the crop specialist will assume that the soils scientist has built and parameterised the soils processes sufficiently for them to get on with the important part of growing crops.

5.2 How hard do we need to work at useability and transparency?

Models such as APSIM are complex, and as such require specialist support, and a range of skills close at hand to support simulation building. A soil scientist will no doubt need crop physiology or agronomy expertise, beyond just a good set of crop parameter values, to ensure that water use, dry matter production and maybe yield (assuming a holistic soil scientist) are in the ball park as drivers for soil processes. Modellers often work in an environment where this broad expertise is available, and is essential for the development of useful and reliable systems tools. It would be naive to expect that this level of support is generally available. There may be parallels between modern aircraft that require expert systems to keep them in the air and agricultural systems models!

To help alleviate this complexity, APSRU has developed a number of derivative tools – decision support products that have aimed to take a subset of systems processes and present them as tools useable by farmers and their advisers. One such product is *Howwet?* Described by Freebairn *et al,* [1996] (www.apsru.gov.au). *Howwet?* estimates soil water accumulation and nitrogen mineralisation during fallow periods, using farmers daily rainfall records and a minimum number of regional soil and climate input parameters. The user selects a location and soil type specified only as a local name. A simple database provides best bet parameter values for a water balance model. The interface was designed to be accessible to non experts, and allows for some calculations to estimate expected crop yield and nitrogen fertiliser requirements.

Howwet? was initially created to:

- Demystify models in the fledgling APSRU years (1991).
- Provide an awareness of models and entry point into modelling.
- Provide added value to rainfall records that farmers routinely collected.
- Provide a decision support and educational tool.

This approach has stood the test of time, with it being routinely used by farmers and consultants. Howwet?'s key features are that it is useful, relatively simple, taking only a few minutes to demonstrate, and it's soil processes are transparent with a range of graphical and tabular presentations available. In recent years, APSRU has strived to reproduce this experience with two more derived products: Whopper Cropper and Yield Prophet

A question arises from this wrapping of complexity: do we put enough effort into making the inner workings of models more accessible, and do we explore some of the many interactions routinely? One observation is that we often take the basic climate record as being correct, a dangerous assumption that can be remedied by a quick graphical scan of the climate record time series. Therefore, the mechanics of being able to scan large data sets quickly needs to be efficient and convenient.

A second question is how much effort goes into building large systems models vs. single issue tools that can be brought out when needed? These questions are not easy to answer and are still being debated in APSRU.

5.3 Access and commercialisation

In recent years, APSRU has sought commercial partners to aid in delivery of model capabilities and tools to farmers and advisers. The Birchip Cropping Group (BCG) and NutrientMS, jointly with APSRU, are responsible for developing and deploying tools and outputs, based on APSIM, directly to consultants and growers. These companies are experienced in the distribution and marketing of agronomic products and services to the rural sector. As well as providing a revenue stream, this separation of model delivery from development frees up time of modellers and scientists for further research and development of APSIM.

In addition to developing commercial licensing agreements, general access to APSIM has changed significantly in recent years. As APSRU moves into a more focussed delivery phase, APSIM's access policy has also changed to be more open and to make APSIM available to all interested users. For more information see www.apsru.gov.au. Whilst APSRU does charge a fee for APSIM to cover ongoing development, it does provide full source code for all modules on www.apsim.info. This provides a degree of transparency for APSIM modules that helps alleviate user's concerns that they are using a 'black box'.

While APSRU has sought to deliver its technology predominantly by the commercial route, we also recognise the importance of contributing to the public good. This is achieved through developing good science, often in collaborative arrangements with other research groups, our contribution to education, and by ultimately delivering a product that will assist the operation of profitable farm and rural enterprises and contribute to a better environment.

6. CONCLUSION

While APSIM is primarily aimed at researchers, an increasing number of derived products have been developed. Adoption by commercial partners is also increasing and it is through these arrangements that consultants and growers who have no prior modelling experience can evaluate a large range of alternative crop and fallow management options. A large international breeding company is also using APSIM to evaluate alternative breeding strategies for maize. The fact that APSIM is in high demand today demonstrates that it is relevant, useful, and stable. These are all key indicators that APSRU's development strategy is sound.

APSRU's experience in developing APSIM, like the process outlined in this paper, has been evolutionary. While APSIM has largely achieved the stated purposes in section 2.1, this has required a substantial investment in the underpinning, strategic infrastructure for model development (estimated costs so far exceed US $ 15 million) that was not without challenges. Not only was it necessary to secure the necessary funds on an on-going basis, it also caused frustration when development did not proceed at the desired speed. Although such tensions will always exist, the approach has evolved into a well designed, structured, and disciplined process. The APSRU group now numbers approximately 80 individuals, most of them contributors to APSIM in some way. It has become a core technology to investigate current and future issues in agricultural and environmental sciences and practice, while providing a vast depository of knowledge, bridging several disciplinary divides. APSIM applications have contributed demonstrably to change in farming practices, particularly in Australia. Given the rapid changes that are currently taking place in rural industries (driven by economic as well as environmental factors such as climate change), the importance of APSIM as a quantitative, predictive tool for scenario development and evaluation is likely to increase.

7. REFERENCES

Brown, P., Huth, N., Banks, P., and Singleton, G., Relationship between abundance of pest rodents and damage to agricultural crops, *Journal of Applied Ecology*, (In Press), 2006.

Donnelly, J.R.; Freer, M.; Salmon, L.; Moore, A.D.; Simpson, R.J.; Dove, H.; Bolger, T.P., Evolution of the GRAZPLAN decision support tools and adoption by the grazing industry in temperate Australia, *Agricultural Systems*, 74(1), 115-139, 2002.

Carberry, P.S., Muchow, R.C., and McCown, R.L., Testing the Ceres-Maize simulation model in a semi-arid tropical environment, *Field Crops Research*, 20, 297-315, 1989.

Chapman, S.C., Hammer, G.L. and Meinke, H., A sunflower simulation model: I. Model development, *Agronomy Journal*, 85, 725-735, 1993.

Dalgliesh, N. & Foale, M. Soil Matters: Monitoring Soil Water and Nitrients in Dryland Farming, CSIRO, Toowoomba, 1998.

Freebairn, D. M., Hamilton, G., Cox, P. G., Glanville, S. F., and Dimes, J. P. HOWWET? Estimating the storage of water in your soil using rainfall records, www.apsru.gov.au, 1996.

Hammer, G.L., Kropff, M.J., Sinclair, T.R. and Porter, J.R., Future contributions of crop modelling - from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement, *European Journal of Agronomy*, 18, 15-31, 2002.

Hammer, G.L. and Muchow, R.C., Quantifying climatic risk to sorghum in Australia's semiarid tropics and subtropics: Model development and simulation, in R.C. Muchow and J.A. Bellamy (eds.), Climatic Risk in Crop Production: Models and Management for the Semiarid Tropics and Subtropics, CAB International, 205-232, 1991.

Hochman, Z., van Rees, H., Carberry, P.S., Holzworth, D., Dalgliesh, N.P., Hunt, J., Brennan, L.E., van Rees, S., Peake, A.S., Poulton, P,L., Huth N.I., and McCown, R.L., Yield Prophet®, an Internet-enabled simulation-based system to assist farmers manage risk in Australia's drought prone environments, *Journal of Experimental Botany*, submitted 2006.

Huth, N. and Holzworth, D., Common Sense in Model Testing, *MODSIM Proceedings*, *ANZMSS*, 2005.

Jakeman, A.J, R.A. Letcher, and J.P. Norton, Ten iterative steps in development and evaluation of environmental models, *Environmental Modelling & Software*, 21(5), 602-614, 2006.

Jones, C.A. and Kiniry, J.R. (Eds.), Ceres-Maize: a simulation model of maize growth and development, *Texas A&M university Press*, *College Station, TX.*, 194, 1986.

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., An overview of APSIM, a model designed for farming systems simulation, *European Journal of Agronomy*, 18:3-4, 267-288, 2003.

Littleboy, M., Silburn, D.M., Freebairn, D.M., Woodruff, D.R., Hammer, G.L., PERFECT - A computer simulation model of Productivity Erosion Runoff Functions to Evaluate Conservation Techniques, *Queensland Department of Primary Industries Bulletin, QB89005*, 1989.

Littleboy, M., Silburn, D.M., Freebairn, D.M., Woodruff, D.R., Hammer, G.L., and Leslie, J.K., Impact of soil erosion on production in cropping systems. I. Development and validation of a simulation model. *Australian Journal of Soil Research*, 30, 757-774, 1992.

Mayer, D.G. and Butler, D.G., Statistical validation. *Ecological Modelling*, 68, 21-32, 1993.

McLean, G. and Doherty, A., REMS – a research experiment management system to support the maintenance and development of crop growth and development models, *Proceedings of the 4th International Crop Science Congress, 26 Sep – 1 Oct 2004, Brisbane, Australia. Published on CDROM*, Web site www.cropscience.org.au, 2004.

Meinke, H. and Stone, R.C., Seasonal and inter-annual climate forecasting: the new tool for increasing preparedness to climate variability and change in agricultural planning and operations, *Climatic Change*, 70, 221-253, 2005.

Moore, A. D., Holzworth, D. P., Herrmann, N. I., Huth, N. I., Keating, B. A., and Robertson, M. J., Specification of the CSIRO Modelling Protocol, LWRRDC Project CPI 9, 2001

Muchow, R.C. and Bellamy, J.A., Climatic risk in crop production: models and management for the semi-arid tropics and subtropics, CAB International, Wallingford, UK, 548, 1991.

Nelson, R.A., Holzworth, D.P., Hammer, G.L. and Hayman, P.T., Infusing the use of seasonal climate forecasting into crop management practice in North East Australia using discussion support software. *Agricultural Systems,* **74,** 393-414, 2002.

Oreskes, N., Shrader-Frechette, K. and Belitz, K., Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641-646, 1994.

Parton, W.J., D.S. Schimel, C.V. Cole, D.S. Ojima, Analysis of factors controlling soil organic levels of grasslands in the Great Plains, *Soil Science Society of America Journal,* 51:1173-1179, 1987.

Probert, M.E., Dimes, J.P., Keating, B.A., Dalal, R.C., Strong, W.M., APSIM's water and nitrogen modules and simulation of the dynamics of water and nitrogen in fallow systems, *Agricultural Systems*, 56, 1-28, 1998c.

QCCA, The data drill, Queensland Centre for Climate Applications. (http://www.dnr.qld.gov.au/resourcenet/datadril.html), 1998.

Ritchie, J.T., A model for predicting evaporation from a row crop with incomplete cover, *Water Resources Research*, 8, 1204-1213, 1972.

Robertson, M.J., Carberry, P.S., Huth, N.I., Turpin, J.E., Probert, M.E., Poulton, P.L., Bell, M., Wright, G.C., Yeates, S.J., Brinsmead, R.B., Simulation of growth and development of diverse legume species in APSIM, *Australian Journal of Agricultural Research*, 53, 429-446, 2002.

Robertson, M.J., Watkinson, A. R., Kirkegaard, J. A., Holland, J. F., Potter, T. D., Burton, W., Walton, G. H., Moot, D. J., Wratten, N., Farre, I., Asseng, S., Environmental and genotypic control of time to flowering in canola and Indian mustard: Control of time to flowering, *Australian Journal of Agricultural Research*, 53(7),793-811, 2002.

Rykiel Jr., E.J., Testing ecological models: the meaning of validation, *Ecological Modelling,* 90, 229-244, 1996.

Shorter, R., Lawn R.J., and Hammer G.L., Improving genotypic adaptation in crops - a role for breeders, physiologists and modellers, *Experimental Agriculture*, 27, 155-175, 1991.

Sinclair, T.R., Water and nitrogen limitations in soybean grain production. I. Model development, *Field Crops Research*, 15, 125-141, 1986.

Uehara, G. and Tsuji, G.Y., Progress in crop modelling in the IBSNAT Project, in: R.C. Muchow and J.A. Bellamy (eds.). Climatic risk in crop production: Models and management in the semi-arid tropics and subtropics, CAB International, Wallingford, 143-156, 1991.

van Keulen, H. and Seligman, N.G., Simulation of Water Use, Nitrogen Nutrition, and Growth of a Spring Wheat Crop, *Pudoc Wageningen,* 310, 1987.

van Oosterom, E., Hammer, G.L. and Chapman, S., Can transition to flowering be modelled dynamically from the gene level? *Contributed Paper in New directions for a diverse planet, Proceedings of the 4th International Crop Science Congress, 26 Sep – 1 Oct 2004, Brisbane, Australia. Published on CDROM,* Web site www.cropscience.org.au, 2004.

Wang, E., Robertson, M.R., Hammer, G.L., Carberry, P., Holzworth, D., Hargreaves, J., Huth, N., Chapman, S., Meinke, H., McLean, G., Design and implementation of a generic crop module template in the cropping system model APSIM, *European Journal of Agronomy*, 18, 121-140, 2003.

Williams, J.R., P.T. Dyke and C.A. Jones., EPIC: a model for assessing the effects of erosion on soil productivity, In: *Analysis of Ecological Systems: State-of-the-Art in Ecological Modeling,* Eds. W.K. Laurenroth et al. Elsevier, Amsterdam, 553-572, 1983.

Woodruff, D.R., 'WHEATMAN' a decision support system for wheat management in subtropical Australia, *Australian Journal of Agricultural Research*, 43, 1483-1499,