

Virtual plant improvement via the cross-fertilisation of ideas

N.I. Huth^a and D.P. Holzworth^a

^a *CSIRO Ecosystem Sciences, Toowoomba, Queensland*
Email: neil.huth@csiro.au

There are a range of plant modelling approaches including process-based models, L-systems or functional-structural models, simplified big leaf models, models aimed at exploring the impacts of genes, through to simple model plants that do little more than provide an uptake term in a water balance. Each model is designed to address different problem domains using a diverse range of software and algorithms.

At a glance these models appear to be very different. However, there is similarity between many of the components within these models. For example, the basic abstraction of a plant canopy into individual phytomers can be very similar in models that differ greatly in complexity. Within many plant models there is also a general concept of resource supply and demand, and internal plant competition for these resources. At the same time, there can be complementarity in the strengths and weaknesses of these different modelling approaches. The strengths of one modelling approach are often the weakness of other approaches. Functional-structural plant models have strengths in modelling the light environment and low-level plant processes, but are not necessarily designed for modelling larger scale farming systems problems. Farming systems models have a long history in modelling the water, carbon, nutrient and management aspects of cropping systems but are struggling to deal with important questions arising from manipulation of plant architecture through genetics or planting geometry.

The requirements for plant models are increasing rapidly to deal with the impacts of changing climates on crops and weeds, environmental impacts on plant growth and resource capture, emergent properties of plant genetic traits or to assist in designing production systems to address issues of food security whilst minimizing environmental impact. With such acceleration in model requirements we must learn from the plant breeders who have dramatically accelerated plant improvement and look for ways to apply these principles to improving our virtual plants.

In this paper we compare the processes of traditional and modern plant improvement and relate them via analogy to the development of virtual plants. Ongoing cross-fertilisation of ideas, evolution and evaluation is required and will require two main things. Software methods are needed to allow the exchange of code or algorithms. Just as importantly, increased dialog between researchers is required to allow the lessons from one model domain or paradigm to be transferred to another.

Keywords: *model reuse, model integration*

1. INTRODUCTION

There are a range of plant modelling approaches including process-based models, L-systems or functional-structural models, simplified big leaf models, models aimed at exploring the impacts of genes, through to simple model plants that do little more than provide an uptake term in a water balance. Each model is designed to address a different problem domain using a diverse range of software and algorithms. At a glance these models appear to be very different. However, there is similarity between many of the components within these models. For example, the basic abstraction of a plant canopy into individual phytomers can be very similar in models developed in very different modelling frameworks (e.g. the APSIM 'process based' wheat model (Wang et al., 2002) and the Functional-Structural plant model ADELwheat (Evers et al., 2007). Within many plant models there is also a general concept of resource supply and demand, and internal plant competition for these resources (e.g. APSIM (Brown et al, 2011) vs ADELwheat (Evers et al., 2010)).

At the same time, there can be complementarity in the strengths and weaknesses of these different modelling approaches. The strengths of one modelling approach are often the weakness of other approaches. Functional-structural plant models have strengths in modelling the light environment and low-level plant processes, but are not necessarily designed for modelling larger scale farming systems problems. Farming systems models have a long history in modelling the water, carbon, nutrient and management aspects of cropping systems but are struggling to deal with important questions arising from manipulation of plant architecture or genetics. There is a growing need for the differing modelling approaches to learn from each other in order to improve.

The need for improvement in our plant modelling capacity has accelerated in recent years with increased expectations of them to help tackle the big problems these times. Models are required to capture the impacts of changing climates on crops and weeds (Toh et al., 2011), environmental impacts on plant growth and resource capture (Dunbabin et al., 2011), emergent properties of plant genetic traits (Hammer et al., 2005) or assist in designing production systems to address issues of food security whilst minimizing environmental impact (Keating et al, 2010). This paper aims to encourage greater exchange of ideas and software components between modellers by drawing parallels to plant breeding, a discipline that has greatly benefited from exchange of ideas and technologies among plant breeders.

2. THE ANALOGY OF VIRTUAL PLANT IMPROVEMENT

We are all very familiar with the concept of plant improvement for it has been practiced around the globe for thousands of years. It is now practiced worldwide by individuals such as gardeners and farmers, or by professional plant breeders employed by organizations such as government institutions, universities, crop-specific industry associations or research centres. Much of the earliest development occurred by the domestication of wild plants, a process which most often occurred in the plant's centre of origin. Landraces of particular species formed via selective processes, natural or managed, became well suited to a particular region or environment. Classical plant breeding was then developed to allow deliberate crossing with closely or distantly related individuals to incorporate multiple desirable properties. In many cases, heterosis, otherwise known as hybrid vigour, occurred and the progeny outperformed both of the parent lines. More recently, numerous techniques have been developed to allow the generation of progeny from genera that ordinarily would not be able to reproduce with each other. Genetically modified plants are regularly developed via the incorporation of DNA, sometimes from a very distant organism, to provide that desirable trait. Genetic modification can produce a plant with the desired trait or traits faster than classical breeding because the majority of the plant's genome is not altered. However, there is much concern within the community regarding the impact of genetically modified plants and possible risks of adverse properties in these new organisms. As a result, GM plants undergo stringent testing to ensure the safety of these new organisms. However, it has been shown that with classical breeding techniques the breeder does not know exactly what genes have been introduced to the new cultivars and so some scientists therefore argue that plants produced by classical breeding methods should undergo the same safety testing regime as genetically modified plants.

No matter what methods are employed, the following steps indicate the normal processes involved in plant improvement:

1. Creation of variation
2. Selection

3. Evaluation
4. Release
5. Multiplication
6. Distribution of the new variety

Traits that breeders have tried to incorporate into crop plants in the last 100 years include:

1. Increased magnitude and quality of crop yield
2. Increased tolerance of environmental pressures (salinity, extreme temperature, drought)
3. Resistance to viruses, fungi and bacteria
4. Increased tolerance to insect pests
5. Increased tolerance of herbicides

The text above is a précis of the Wikipedia article on plant breeding (http://en.wikipedia.org/wiki/Plant_breeding, Accessed 26 July 2011). As modellers, we are used to the idea that virtual plants can be used to describe aspects of real plants. Furthermore, many researchers now perform virtual experiments *in silico* to evaluate virtual plants as part of a screening exercise even before such plants have been developed using the techniques described above (Hammer et al., 2005). However, it is not clear that the similarities between historical plant improvement and modern plant model development have been noticed. Whilst the analogy is far from complete, there is enough similarity to allow us to explore the history of virtual plant development, and obtain some lessons on what is required to take us into our own virtual green revolution. For example, the history of plant model development could be expressed using a simple adaptation of the history above.

So, what would a paraphrase of the history of plant improvement look like if we drew analogies between genes and source code, traits and predictive capability, species and model?

The following text is one possible version:

We are all very familiar with the concept of plant model development for it has been practiced around the globe for several decades. It is now practiced worldwide by individuals such as agriculturalists and farmers, or by professional modellers employed by organizations such as government institutions, universities, crop-specific industry associations or research centres. Much of the earliest development occurred by the adaptation of initially small modelling efforts, a process which most often occurred in the model's centre of origin. Versions of particular models formed via ongoing research become well suited to a particular region or environment. Software techniques were then developed to allow integration with closely or distantly related models to incorporate multiple desirable model components. In many cases, improved model functionality occurred and these new 'systems models' provided greater utility than the model components in isolation. More recently, numerous techniques have been developed to allow integration of models that ordinarily would not be able to communicate with each other. Programmatically modified models are regularly developed via the incorporation of logic, sometimes from a very different modelling approach, into an existing model to provide that extra functionality. Such modification can produce a model with the desired functionality faster than traditional processes because the majority of the model's code is not altered. However, there is much concern within the modelling community regarding the validity of modern complex models and possible risks of adverse emergent properties in these new models. As a result, complex plant models undergo stringent testing to ensure the scientific and software quality of these new models. However, it has been shown that with traditional unmanaged coding techniques the model developer does not know exactly what code has been introduced to the new models versions and so some scientists therefore argue that models produced by even traditional coding methods should undergo the same testing regime as modern complex plant models.

No matter what research or software methods are employed, the following steps indicate the normal processes involved in plant model improvement:

1. *Creation of individual plant models for a particular environment*
2. *Testing of models in a wider range of environments*
3. *Collaborative development of model enhancements*
4. *Publication of new model(s)*

5. *Distribution of the new model to user base*

Functionality that model developers have tried to incorporate into plant models in recent decades includes:

1. *Improved predictions of magnitude and quality of crop yield*
2. *Responses to environmental pressures (salinity, extreme temperature, drought)*
3. *Responses to plant pathogens*
4. *Responses to plant pests*
5. *Evolution of herbicide tolerance in weeds*

It is not surprising that the analogy fits so neatly. Both plant improvement and model development have made progress via adaptation of localized work, addition of beneficial elements from similar material, or novel incorporation of properties from something previously unrelated. The continuation of such processes is going to be crucial to the scientific community in ensuring we can reach the goals set for us in a world facing threats to food security. So it is prudent for us to ask ourselves if we have learned the lessons of the history of plant breeding that will assist us in creating our own virtual green revolution.

3. FACILITATING CROSS-FERTILISATION

There are a range of aspects to model development that can facilitate the exchange of ideas or model source code. Here we discuss just four of them.

3.1. Integrative Frameworks

There has been much discussion for the need for software frameworks for the integration of modelling components. Frameworks allow virtual experiments to be conducted. The ability to quickly and easily replace a process or model with another whilst keeping all other processes the same allows for quick comparison and evaluation of ideas. This in turn leads to scientific debate about the merits or otherwise of particular approaches. Before long, communities of like minded developers and modellers tend to coalesce around the framework. The APSIM experience bears this out with several organizations contributing not just to the development of the framework but also to the cross-fertilisation of ideas and approaches. This same phenomenon can be seen with large efforts happening in Europe, the United States and elsewhere.

Exactly what the framework looks like or how it works is, to a large degree, irrelevant provided that it gives the required functionality and facilitates collaboration, peer review and scientific debate. Multsch et al. (2011) give some design characteristics that a framework must exhibit: replaceability, expandability, independence and interactivity. These are all worthwhile attributes inherent in all good frameworks.

There are many examples where frameworks are built and extended by “crossing” approaches with other frameworks. Holzworth et al. (2010) describe how the APSIM development team has adopted many of the reflection metadata tags that are used in TIME (The Invisible Modelling Environment - a software development framework for creating, testing and delivering environmental simulation models (Rahman et al. 2004)). By doing so, they have provided a mechanism where models written for APSIM or TIME are at least partially compatible.

Despite this, within the modelling community there still seems to be a desire to build Yet Another Modelling Framework (YAMF). van Evert et al (2005) discusses this in some length and suggest that these efforts do have some merit but only if new efforts capitalize on the work of others. Occasionally these new works will produce a desirable trait that brings something new to the modelling community. These ideas can then be crossed with existing frameworks to build a better platform. This is evolution at work.

3.2. Simpler component interfaces

The provision of simpler component interfaces provides an alternative to the adoption of a single overarching integrative framework. In some cases, the simple wrapping of a piece of model functionality into a component with a simple, but well defined interface may be all that is required. There are now modern software techniques that reduce the amount of plumbing code (e.g. the code to get and set variable values) and allow the model builder to focus on the problem domain (Holzworth et al. 2010, Rahman et al. 2004). They are in essence, framework independent. These allow the framework to use common software techniques, such as reflection or introspection, to discover variables, states, parameters and outputs at runtime without them having to be explicitly introduced to the framework. By coding a model or process this way,

any framework should be able to use the code without much, if any, modification. The ideal would be an accessible library of models, processes, functions that can be used in a wide range of frameworks; something akin to the R project (<http://www.r-project.org/>). However, underpinning such efforts must be a willingness by developers at the outset to develop models suitable for use by others, and a desire by others to look for existing components for rapid deployment before reinventing the wheel.

3.3. Better technologies for describing integrated modelling

Integration of model concepts requires better ways to describe the individual components, their interfaces, their data and how they are to be incorporated into a new single entity. Many of these issues are not particular to plant modelling alone and solutions have been developed within the wider software community. One particular technology that exemplifies this is the extensible markup language (XML) which is now widely used in the plant modelling community (Holzworth and Huth, 2009). Other developers have adopted or adapted Python, ontologies, UML, or declarative programming techniques to help describe the methods or data used in model integration (Multsch et al. 2011, van Evert et al., 2005). Once again, however, such efforts will not produce fruit unless there is a willingness within the modelling community to take advantage of them.

3.4. Sharing of ideas

The three examples above are all software solutions to the problems arising from model improvement via ongoing integration of ideas. All of them are extensively utilized within the modelling community and the number of publications on each of these continues to grow each year. No doubt, future developments in programming languages, software design techniques, logic expression and computing in general will continue to find even better ways to integrate modelling components. However, though we are software developers we shouldn't forget that we are also researchers, and people.

Some of the best ways to integrate, evaluate, share and distribute ideas are via writing and speaking, and doing so in a broad forum, rather than just by ourselves or with likeminded individuals. Ongoing incremental improvement of aspects of our own thinking without input from a broader school of thought may provide an improvement in our models. However, it is likely that such progress, whilst beneficial in itself, will take us about as far as the domestication of wild plants, a process which took several thousand years. Integration of quite similar model components, different in parentage but similar in scale and scope has accelerated plant model development in much the same way that conventional plant breeding has increased crop yield since the time of Gregor Mendel. However, we argue that for even more rapid progress, quantum leaps like those often promised by genetic modification, will require incorporation of material from currently distant relatives in the virtual plant realm.

To do this, improved communications amongst a wider range of plant modellers will be required. Even a superficial analysis of the scientific literature suggests that references to significantly different modelling approaches are not common within the plant modelling literature. Furthermore, when reference is made to other modelling work the reference is often brief and superficial. A survey of citations of the key APSIM modelling papers (Keating et al., 2003, Wang et al., 2002) showed that in most cases APSIM was only cited as an example of another system for modelling farming systems. Descriptions of the important aspects of the model, and any ideas gained from it were not often included. The authors confess that this is also an obvious deficiency in their own literature when citing other models and so we will have to improve our own efforts in this regard.

4. DISCUSSION AND CONCLUSIONS

We suggest there is opportunity for modellers in different modelling groups and different modelling domains to learn more from each other than is currently the case. Much work has been done on software methods to allow the exchange of code or algorithms. However, increased dialog between researchers is required to allow the lessons from one domain, or modelling paradigm, to be transferred to another. Conferences such as MODSIM provide a venue for such communications and we encourage those attending to seek out those experienced in different modelling approaches and work on the cross-fertilisation of ideas.

ACKNOWLEDGMENTS

The authors would like to acknowledge the many researchers with whom they have exchanged ideas within the Agricultural Production Systems Research Unit (APSRU) and the APSIM Initiative (AI).

REFERENCES

- Brown, H.E., N. Huth, and D. Holzworth (2011). A potato model built using the APSIM Plant.NET Framework. In Chan, F., D. Marinova and R.S. Anderssen (eds) MODSIM2011 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011.
- Dunbabin, V.M., M. Airey, A.J. Diggle, M. Renton, Z. Rengel, R. Armstrong, Y. Chen, and K.H.M. Siddique (2011). Simulating the interaction between plant roots, soil water and nutrient flows, and barriers and objects in soil using ROOTMAP. In Chan, F., D. Marinova and R.S. Anderssen (eds) MODSIM2011 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011.
- Evers, J.B., J. Vos, C. Fournier, B. Andrieu, M. Chelle, and P.C. Struik (2007). An architectural model of spring wheat: Evaluation of the effects of population density and shading on model parameterization and performance. *Ecological Modelling*, 200, 308-320.
- Evers, J.B., J. Vos, X. Yin, P. Romero, P.E.L. van der Putten, and P.C. Struik (2010). Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *Journal of Experimental Botany*, 61, 2203-2216.
- Hammer, G.L., S. Chapman, E. van Oosterom, and D.W. Podlich (2005). Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. *Australian Journal of Agricultural Research*, 56, 947-960.
- Holzworth, D., and N. Huth (2009). Reflection + XML Simplifies Development of the APSIM Generic PLANT Model. In: Anderssen, R.S., Braddock, R.D., Newham, L.T.H. (Eds.), 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia, pp. 887-893.
- Holzworth, D.P., N.I. Huth, and P.G. de Voil (2010). Simplifying environmental model reuse. *Environmental Modelling and Software* 25:269-275.
- Keating, B.A., P.S. Carberry, P.S. Bindraban, S. Asseng, H. Meinke, and J. Dixon (2010). Eco-efficient agriculture: concepts, challenges, and opportunities. *Crop Science*, 50 (supplement1), s-109-s-119.
- Keating, B.A., P.S. Carberry, G.L. Hammer, M.E. Probert, M.J. Robertson, D. Holzworth, N.I. Huth, J.N.G. Hargreaves, H. Meinke, Z. Hochman, G. McLean, K. Verburg, V. Snow, J.P. Dimes, M. Silburn, E. Wang, S. Brown, K.L. Bristow, S. Asseng, S. Chapman, R.L. McCown, D.M. Freebairn, and C.J. Smith (2003). An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18, 267-288.
- Multsch, S., P. Kraft, H.-G. Frede, and L. Breuer (2011). Development and application of the generic plant growth modeling framework (PMF). In Chan, F., D. Marinova and R.S. Anderssen (eds) MODSIM2011 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011.
- Rahman, J.M., S.P. Seaton, and S.M. Cuddy (2004). Making frameworks more useable: using model introspection and metadata to develop model processing tools. *Environmental Modelling and Software*, 19, 275-284.
- Toh, R., J. Hanan, K. Dhileepan, R.G. Shivas and S.W. Adkins (2011). Simulation of parthenium weed canopy under changing climate using L-systems. In Chan, F., D. Marinova and R.S. Anderssen (eds) MODSIM2011 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2011.
- van Evert, F., D. Holzworth, R. Muetzelfeldt, A. Rizzoli, and F. Villa (2005). Convergence in integrated modeling frameworks. In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 745-750. ISBN: 0-9758400-2-9.
- Wang, E., M.J. Robertson, G.L. Hammer, P.S. Carberry, D. Holzworth, H. Meinke, S.C. Chapman, J.N.G. Hargreaves, N.I. Huth, and G. McLean (2002). Development of a generic crop model template in the cropping system model APSIM. *European Journal of Agronomy*, 18, 121-140.