Re-inventing model-based decision support with Australian dryland farmers. 3. Relevance of APSIM to commercial crops

P. S. Carberry\textsuperscript{A,F}, Z. Hochman\textsuperscript{B}, J. R. Hunt\textsuperscript{C}, N. P. Dalgliesh\textsuperscript{A}, R. L. McCown\textsuperscript{A}, J. P. M. Whish\textsuperscript{A}, M. J. Robertson\textsuperscript{D}, M. A. Foale\textsuperscript{B}, P. L. Poulton\textsuperscript{A}, and H. van Rees\textsuperscript{C,E}

\textsuperscript{A}Agricultural Production Systems Research Unit (APSRU), CSIRO Sustainable Ecosystems, PO Box 102, Toowoomba, Qld 4350, Australia.
\textsuperscript{B}Agricultural Production Systems Research Unit (APSRU), CSIRO Sustainable Ecosystems, 306 Carmody Road, St Lucia, Qld 4067, Australia.
\textsuperscript{C}Birchip Cropping Group, PO Box 85, Birchip, Vic. 3483, Australia.
\textsuperscript{D}CSIRO Sustainable Ecosystems, Private Bag 5, PO Wembley, WA 6913, Australia.
\textsuperscript{E}Cropfacts P/L, 69 Rooney Rd, RSD Strathfieldsaye, Vic. 3551, Australia.
\textsuperscript{F}Corresponding author. Email: peter.carberry@csiro.au

Abstract. Crop simulation models relevant to real-world agriculture have been a rationale for model development over many years. However, as crop models are generally developed and tested against experimental data and with large systematic gaps often reported between experimental and farmer yields, the relevance of simulated yields to the commercial yields of field crops may be questioned. This is the third paper in a series which describes a substantial effort to deliver model-based decision support to Australian farmers. First, the performance of the cropping systems simulator, APSIM, in simulating commercial crop yields is reported across a range of field crops and agricultural regions. Second, how APSIM is used in gaining farmer credibility for their planning and decision making is described using actual case studies.

Information was collated on APSIM performance in simulating the yields of over 700 commercial crops of barley, canola, chickpea, cotton, maize, mungbean, sorghum, sugarcane, and wheat monitored over the period 1992 to 2007 in all cropping regions of Australia. This evidence indicated that APSIM can predict the performance of commercial crops at a level close to that reported for its performance against experimental yields. Importantly, an essential requirement for simulating commercial yields across the Australian dryland cropping regions is to accurately describe the resources available to the crop being simulated, particularly soil water and nitrogen.

Five case studies of using APSIM with farmers are described in order to demonstrate how model credibility was gained in the context of each circumstance. The proposed process for creating mutual understanding and credibility involved dealing with immediate questions of the involved farmers, contextualising the simulations to the specific situation in question, providing simulation outputs in an iterative process, and together reviewing the ensuing seasonal results against provided simulations.

This paper is distinct from many other reports testing the performance and utility of cropping systems models. Here, the measured yields are from commercial crops not experimental plots and the described applications were from real-life situations identified by farmers. A key conclusion, from 17 years of effort, is the proven ability of APSIM to simulate yields from commercial crops provided soil properties are well characterised. Thus, the ambition of models being relevant to real-world agriculture is indeed attainable, at least in situations where biotic stresses are manageable.

Additional keywords: APSIM, crop simulation model, validation, decision support systems, commercial yield.

Introduction
The goal of crop simulation models being relevant to real-world agriculture has been the rationale for model development over many years (Muchow and Bellamy 1991; Ritchie 1991; McCown et al. 2002; Jones et al. 2003; Stöckle et al. 2003; van Ittersum and Donatelli 2003; van Ittersum et al. 2003; van Keulen 2007). In Australia, Carberry et al. (2002) described the FARMSCAPE (Farmers’, Advisers’, Researchers’ Monitoring, Simulation, Communication and Performance Evaluation) experience, which questioned whether farmers could value crop simulation as decision support and, if so, how such models could be applied cost-effectively. In addressing the goal of models used in farm management, Carberry \textit{et al.} (2002) point to 3 requirements that are likely needed. First, the proposed model must be able to represent yields of actual commercial crops as tested through comparison of simulated and paddock yields. Second, the model needs to be flexible and comprehensive in its capability to address issues of interest to farmers and their advisers and suppliers.
Lastly, systems are required to deliver the models to industry cost-effectively. In further exploring the FARMSCAPE experience in this series of papers introduced by McCown et al. (2009, this issue), this particular paper addresses (a) how well the cropping systems simulator, APSIM (McCown et al. 1996; Keating et al. 2003), simulates commercial crop yields and (b) how APSIM is used to assist farmers’ planning and decision making in the dryland cropping regions of Australia. Other papers in the series address issues such as the soil information required for paddock-specific simulation (Dalgliesh et al. 2009), using the internet to deliver simulations (Hochman et al. 2009b) and an interpretation of the FARMSCAPE experience using social theory (McCown et al. 2009).

In Australia, the French-Schultz model for water-use efficiency (WUE) is widely used by researchers and industry to establish potential crop yields (French and Schultz 1984a, 1984b; Angus and van Herwaarden 2001). Its many applications to Australian wheat crops largely confirm the view that most commercial crops fall below their water-use potential in Australia (Cornish and Murray 1989; Beeston et al. 2005; Sadras and Angus 2006; Hochman et al. 2009a). Typical causes for lower than expected yields are nitrogen deficiency, biotic stresses caused by pests, disease, and weeds (Angus and van Herwaarden 2001) or harvest losses (Robertson et al. 2000), although Sadras and Angus (2006) point towards low phosphorus availability, late sowing, and subsoil chemical constraints as agronomic factors that also constrain yields in Australia. With a lack of attention to good agronomy supposedly remaining as a major limiting factor to production in Australia (Freebairn et al. 2005), the question remains on the likelihood of crop models, which do not explicitly account for biotic stresses or soil chemical constraints, being relevant to commercial crops grown under such constraints.

The observed yield gap between farm and experimental or modelled yields (Davidson 1962; Sadras and Angus 2006) is in contrast to the claims by Carberry et al. (2002) that simulated yields from APSIM tested well against commercial yields obtained on farms. While Robertson et al. (2000) demonstrated that APSIM closely simulated mungbean yields measured in commercial paddocks in Australia, simulations over-predicted actual harvested yield due to significant machine-harvest losses. However, more recently, APSIM has been shown to simulate yields as measured in commercial paddocks in Australia, explaining 74% of variation in wheat yields (Sadras et al. 2003), 85% of the variation in sorghum yields (Whish et al. 2005), and 61% of chickpea yield variability (Whish et al. 2007). In similar tests in Argentina, Merceau et al. (2007) found that a locally calibrated CROPGRO-Soybean model explained only 41% of the measured variation in yields of soybean grown on commercial farms.

It might be expected that model accuracy is primarily of concern to scientists and utility is primarily of concern to farmers. However, in our experience, the fundamental scepticism of farmers concerning the ability of a theory-based simulator to mimic their reality makes farmers harsh judges of simulator performance. On the other hand, when practical intervention in management is the scientists’ aim, simulator accuracy as the performance criterion is as important as utility to managers. This paper looks at the performance of APSIM as the key scientific tool used in FARMSCAPE for intervening in farmers’ planning and decision processes. Whereas the ability of APSIM to simulate yields of a wide range of crops has been previously tested using experimental data (Keating et al. 2003), the evaluation in this paper moves from ideal scientific conditions to conditions of real-world practice. This entails both empirical comparisons with measured farm paddock yields and the experiences of farmers using the simulator in dealing with management uncertainty. Adequate evaluation of the simulator in FARMSCAPE intervention in practice necessarily entails both of these ‘hard’ and ‘soft’ aspects.

‘Hard’ evaluation of APSIM: testing against commercial crops

Description of available data

Accessible accounts of the performance of APSIM in simulating commercial crops in Australia are collated into Table 1. These accounts range from publications in refereed journals to data contained in conference or industry publications and unpublished data. In this collection, consisting of crops of barley, canola, chickpea, cotton, maize, mungbean, sorghum, sugarcane, and wheat monitored over the period 1992–2007, APSIM simulated grain, cane, and fibre yields obtained from over 700 commercial crops grown under a wide range of environmental and agronomic conditions from all cropping regions of Australia.

Details on each dataset and simulation are contained in the references cited in Table 1. Briefly, the reported yields for the crops in each study were measured by one of several procedures, including multiple hand-harvested plots randomly collected within a crop, large plots harvested by machine and measured using a weigh bin, on-board yield monitor equipment fitted to commercial harvesters, and yields calculated from total paddock production divided by harvested area. Each measurement procedure contains some likely source of error, either over-estimating commercial yield from small plot harvests, due to lower-than-commercial harvesting loss, or possible calibration errors with yield monitoring equipment. However, such procedures for reporting commercial yields are accepted within the grains industry of Australia and so are appropriate for assessing whether a remote yield estimate (via simulation) corresponds to the reported yields of commercial crops. For all crops represented in Table 1, APSIM simulations were run with actual rainfall data for the cropping season collected from sites close to the crop of interest, usually at the farmer’s house. Many simulations used rainfall, temperature, and radiation data collected from automatic weather stations situated next to the monitored crop. Where temperature and radiation data were unavailable from the farm, data were obtained from the nearest official climate station using archived Australian climate data (Jeffrey et al. 2001).

Actual crop agronomic information on sown variety, sowing date, density, and row spacing were also used. In most cases, input parameters to describe the soil were as measured in-situ in the paddock where each crop was grown. At each site, plant-available water capacity (PAWC), defined as the difference in volumetric water content between drained upper limit (DUL) and crop lower limit (CLL), was either explicitly
Table 1. The performance of APSIM in simulating commercial crops in Australia and, where available, the referenced source

<table>
<thead>
<tr>
<th>Crop</th>
<th>Region</th>
<th>Details</th>
<th>n</th>
<th>R²</th>
<th>Slope</th>
<th>Intercept</th>
<th>RMSD (t/ha)</th>
<th>Range</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barley</td>
<td>South</td>
<td>Sowing date, variety</td>
<td>45</td>
<td>0.85</td>
<td>0.87</td>
<td>0.06</td>
<td>0.53</td>
<td>0.0–4.4</td>
<td>Z. Hochman et al., unpubl. data</td>
</tr>
<tr>
<td>Canola</td>
<td>South</td>
<td>Sowing date, variety</td>
<td>42</td>
<td>0.86</td>
<td>0.77</td>
<td>0.56</td>
<td>0.48</td>
<td>0.4–5.5</td>
<td>Robertson and Kirkegaard (2005)</td>
</tr>
<tr>
<td>Chickpea</td>
<td>North</td>
<td>Sowing date, density, row spacing</td>
<td>37</td>
<td>0.61</td>
<td>0.79</td>
<td>0.27</td>
<td>0.30</td>
<td>0.5–2.8</td>
<td>Whish et al. (2007)</td>
</tr>
<tr>
<td>Cotton¹</td>
<td>North</td>
<td>Sowing date, row config.</td>
<td>30</td>
<td>0.70</td>
<td>1.00</td>
<td>0.49ᵇ</td>
<td>1.37ᵇᶜ</td>
<td>1.2–8.2ᵇ</td>
<td>Bange et al. (2005)</td>
</tr>
<tr>
<td>Maize</td>
<td>North</td>
<td>Sowing date, variety, row config.</td>
<td>13</td>
<td>0.93</td>
<td>1.03</td>
<td>−0.24</td>
<td>0.66</td>
<td>2.0–9.6</td>
<td>Robertson et al. (2003)</td>
</tr>
<tr>
<td>Mungbean</td>
<td>North</td>
<td>Sowing date, variety</td>
<td>23</td>
<td>0.73ᶜ</td>
<td>0.75ᶜ</td>
<td>0.29ᶜ</td>
<td>0.28</td>
<td>0.3–2.3</td>
<td>Robertson et al. (2000)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>North</td>
<td>Sowing date</td>
<td>19</td>
<td>0.98</td>
<td>0.94</td>
<td>0.19</td>
<td>0.35</td>
<td>0.0–8.2</td>
<td>Foale and Carberry (1996)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>North</td>
<td>Sowing date, row config.</td>
<td>29</td>
<td>0.85</td>
<td>0.95</td>
<td>0.32</td>
<td>0.41ᶜ</td>
<td>2.0–6.0</td>
<td>Whish et al. (2005)</td>
</tr>
<tr>
<td>Sugarcane</td>
<td>North</td>
<td>Plant/ratoon, irrig./dryland, variety</td>
<td>35</td>
<td>0.72</td>
<td>0.80</td>
<td>32.2ᵈ</td>
<td>1.94ᵈ</td>
<td>5.0–300.0ᵈ</td>
<td>Keating et al. (1999)</td>
</tr>
<tr>
<td>Wheat</td>
<td>South</td>
<td>Sowing date, variety, soil</td>
<td>55</td>
<td>0.74</td>
<td>0.98</td>
<td>0.01</td>
<td>0.19</td>
<td>0.7–3.8</td>
<td>Sadras et al. (2003)</td>
</tr>
<tr>
<td>Wheat</td>
<td>South</td>
<td>Sowing date, soil</td>
<td>28</td>
<td>0.72</td>
<td>0.89</td>
<td>0.37</td>
<td>0.43</td>
<td>1.1–4.9</td>
<td>Unpubl. data</td>
</tr>
<tr>
<td>Wheat</td>
<td>National</td>
<td>Sowing date, variety, soil</td>
<td>334</td>
<td>0.71</td>
<td>0.80</td>
<td>0.45</td>
<td>0.80</td>
<td>0.0–7.8</td>
<td>Hochman et al. (2009a)</td>
</tr>
<tr>
<td>Wheat</td>
<td>West</td>
<td>Sowing date, variety</td>
<td>7.6</td>
<td>0.52</td>
<td>0.74</td>
<td>0.13</td>
<td>0.55</td>
<td>0.6–2.3</td>
<td>Weeks et al. (2007)</td>
</tr>
<tr>
<td>Wheat</td>
<td>West</td>
<td>Soil type</td>
<td>65</td>
<td>0.80ᶠ</td>
<td>0.80ᶠ</td>
<td>0.45ᶜ</td>
<td>0.30</td>
<td>0.4–3.7</td>
<td>Oliver et al. (2006)</td>
</tr>
</tbody>
</table>

ᵃSimulations by the OZCOT cotton model that is included in the APSIM systems model to simulate cotton crops.
ᵇBales/ha (1 bale = 227 kg).
ᶜEstimated from digitised data taken from the cited reference.
ᵈt cane freshweight/ha.

measured using the techniques described by Dalgliesh and Foale (1998) or it was estimated from the wettest and driest soil water values measured under the crop of interest during the growing season. Measurements of bulk density, soil organic matter, pH, and electrical conductivity (EC) were undertaken during the procedure to determine PAWC. In relatively few cases a PAWC value was substituted using a similar soil type represented in a soil database (Dalgliesh et al. 2009, this issue). At all sites, prescribing soil samples were taken to establish initial conditions for soil water content and nitrate-N concentration.

In summary, most of the simulations reported in Table 1 were initialised with agronomic and soil data collected from the crop and site being simulated.

Overall performance

Across the 14 case studies, APSIM accounted for 52–98% of measured variation in paddock yields. The validations demonstrated some tendency to over-predict low yields and under-predict high yields.

The validation exercise by Hunt et al. (2006), using a subset of the data reported by Hochman et al. (2009a) (Table 1), is enlightening. Hunt et al. (2006) showed that simulated wheat yields improved for the subset of paddocks which had been characterised for soil properties, particularly PAWC: across 64 crops, 68% of simulated results were within 0.5 t/ha of measured yields and the prediction $R^2$ was 0.68. In contrast, the remaining paddocks from the 173 total had soil type assigned to them based on local knowledge and, consequently, only 49% of simulated results were within 0.5 t/ha and the $R^2$ equalled 0.54. Clearly, a strong lesson from this case is the importance of localised soil characterisation to provide APSIM with the opportunity to closely simulate actual paddock yield. Sadras et al. (2003) also found that predicted yields dramatically improved when field-determined soil water properties were used instead of estimates based on soil texture.

The performance of APSIM across this range of case studies provides evidence for two assertions. Firstly, APSIM is relevant for simulating commercial yields and, secondly, the key to such is providing the model with contextualised data for the simulation being undertaken. The supplies of water and nitrogen resources from the soil to the crop are the key determinants of yield in Australian agriculture and they are also the key variables needed by APSIM to estimate paddock yield. Simulation accuracy is therefore largely dependent on the quality of the data describing the soil resources.

‘Soft’ evaluation of APSIM: usefulness to farmers

Realistic simulation of commercial crop yields over many seasons and situations created a strong validation case for APSIM. However, this is not the sole criterion to establish credibility for simulation among farmers and their advisers. APSIM deployment in addressing real-life questions, providing outputs which either aligned with farmer expectations or motivated new thinking, also created credibility. In such APSIM applications, 4 categories of simulation have been identified (Carberry et al. 2002, Hochman et al. 2002):

1. benchmarking, where the performance of a field crop is compared with the simulation output of APSIM;
2. tactical planning, where agronomic management options for the current season are evaluated based on the known status of the system early in the season;
(3) yield forecasting, where final yield probabilities are forecast at any time before or during a season;
(4) scenario exploration, where alternative management options are explored at a strategic level.

Since 1992 when FARMSCAPE commenced, APSIM has been used in many hundreds of face-to-face interactions with farmers, which encompass these 4 categories and create and enhance its credibility to address their questions of interest. With the introduction of the Yield Prophet® forecasting system (www.yieldprophet.com.au), whereby APSIM is now available for online internet applications (Hunt et al. 2006; Hochman et al. 2009b, this issue), there have been multiple thousands of APSIM simulations run specifically for farmers. Most of these applications have remained unreported. Here, for the purpose of describing and understanding the types of credibility-forming interactions undertaken, examples are provided for each of the 4 application types. These examples are not necessarily the best nor most influential of their type, but they provide the context of an actual situation, they are purposely selected from the different agro-ecological cropping zones around Australia, and they describe both how APSIM was set up and then run with the participating farmers.

**Benchmarking**

Carberry et al. (2002) described in some detail the initial meeting with farmers in 1993 whereby a cotton crop was benchmarked against its simulated yield, a process which led to a session of subsequent ‘what-if?’ analysis and discussion (WifAD). For much of the early FARMSCAPE interactions, this same process was followed: initial benchmarking leading to subsequent questions raised by participating farmers. In another simple example of this benchmarking process (example #1), the performance of a sorghum crop grown in south-eastern Queensland was assessed against APSIM simulations (Table 2).

The crop was harvested on 20 February 1996, yielding 6.4 t/ha (12% moisture). Given information about the crop’s management, the climate and the starting water and N conditions, APSIM also predicted a yield of 6.4 t/ha. This result suggested that the sorghum crop achieved its yield potential for the 1995–96 season given the amount of supplied N, and reveals that all sorghum crop achieved its yield potential for the 1995–96 season.

**Tactical planning**

In 2002, a research effort was initiated to introduce tools such as soil characterisation, deep soil coring, and APSIM simulations to farmers operating in the northern wheatbelt of Western Australia. Prior to this intervention, such tools were not generally exposed to local farmers. Likewise, while APSIM had been extensively tested for simulation of experimental crop performance in this environment (Asseng et al. 1998), its application on farms had yet to be trialled. This cropping region of Australia is distinct because of its low latitude and thus, relatively warm winter and hot spring and its predominantly winter rainfall. The annual average rainfall ranges from over 500 mm near the Indian Ocean to the west down to 300 mm in the north-eastern wheatbelt. The soils are highly variable in texture, with the favoured cropping soil, a deep yellow sand, common across the whole region but intermingled with soils of higher and lower clay content.

In the examples below, APSIM was tested by local farmers in exploring possible tactical management options in the 2002 winter cropping season (May–December). Accordingly, soils were characterised for PAWC prior to the season, and starting soil water and nitrogen status were measured. A series of meetings was held with the participating farmers as part of their local farmer groups. What are reported here are the simulations provided to the farmers early in the season in response to their tactical management enquiries and again at the season’s end to assess APSIM simulations against actual crop performance.

In example #2, APSIM simulations were undertaken for a farm in the northern wheatbelt in Western Australia making use of 102 years of climate data (1900–2001). The farm was situated in a favourable environment for wheat production on its deep yellow sand soil type (Table 2). APSIM simulations were undertaken on 25 July 2002 and reported to the farmer and his neighbours on this date. Of particular interest to the farmers at this time was whether further N fertiliser should be topdressed on their crops given the relatively poor start to the 2002 season.

**Table 2. Details of the 5 example applications of APSIM (further explanation is in the text)**

<table>
<thead>
<tr>
<th>Example</th>
<th>Location</th>
<th>Crop</th>
<th>Sowing date and rate</th>
<th>Available soil water (mm)</th>
<th>Soil nitrate-N (kg N/ha)</th>
<th>Fertiliser rate (kg N/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Dalby, Qld; 27.2°S, 151.3°E; 344 m</td>
<td>Sorghum cv. Buster</td>
<td>14 Oct. 1995, 7 plants/m²</td>
<td>146</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>#2</td>
<td>Mingnew, WA; 29.1°S, 115.3°E; 153 m</td>
<td>Wheat cv. Brookton</td>
<td>22 May 2002, 150 plants/m²</td>
<td>34</td>
<td>94</td>
<td>9</td>
</tr>
<tr>
<td>#3</td>
<td>Dalwallinu, WA; 30.3°S, 116.7°E; 343 m</td>
<td>Wheat cv. Arrino</td>
<td>24 June 2002, 100 plants/m²</td>
<td>5</td>
<td>62</td>
<td>6</td>
</tr>
<tr>
<td>#4</td>
<td>Ardrossan, SA; 34.4°S, 137.9°E; 22 m</td>
<td>Wheat cv. AGT Scythe</td>
<td>13 May 2006, 150 plants/m²</td>
<td>–7</td>
<td>105</td>
<td>51</td>
</tr>
<tr>
<td>#5</td>
<td>Dalby, Qld; 27.2°S, 151.3°E; 344 m</td>
<td>Cotton</td>
<td>15 Oct. 1997, 25%, 50%, 75% PAWC</td>
<td>Non-limiting</td>
<td>Non-limiting</td>
<td>0</td>
</tr>
</tbody>
</table>
and a consistently negative SOI phase (Stone and Auliciems 1992) at the time of the simulation meeting.

For each year between 1900 and 2001, APSIM simulated the yield of wheat using the known agronomic information (Table 2), the measured soil water and N values, and the local rainfall from the date of initiation of the simulation (15 April) until the date the simulation was run (25 July), with the historical climate record substituted thereafter. Consequently, Fig. 1a presents the probability of exceedence of simulated grain yields for the featured crop grown at Mingenew. As of 25 July 2002, simulated yields across all years ranged from 0.2 to 6.0 t/ha, with a median yield of 2.1 t/ha and a 30% chance of achieving greater than 2.5 t/ha. The output specified for the starting conditions in 2002 differed little from the simulated long-term yield distribution for a hypothetical continuous wheat scenario with the same agronomic inputs but no specified starting soil or rainfall conditions (data not shown). In other words, the initial 2 months of crop growth in the 2002 season had little influence on changing the yield outlook from the long-term probabilities. However, Fig. 1a also shows the probability of exceedence for the subset of years where the SOI phase was in a consistently negative state during the May–June period, as was the case in July 2002. In such cases, the simulated median yield decreased to 1.7 t/ha with less than 10% chance of achieving 2.5 t/ha. When 23 kg N/ha was added as topdressed fertiliser on 25 July each year of the simulations, the simulated yield response was negligible in most years, with some seasons producing increased yields but many seasons showing yield reductions (Fig. 1b). This was likely to be due to the added fertiliser enhancing vegetative growth and increasing early soil water use at the cost of later grain production.

In this case the participating farmer applied no topdressed N fertiliser to this wheat crop, in contrast to his normal practice. The strong influence on this decision was the discovery of a large quantity of mineral N in the soil before sowing. As it happened, the 2002 season had 252 mm total in-crop rainfall, with few falls greater than 10 mm per day. The crop yielded 2.6 t/ha while APSIM simulated a final yield of 3.3 t/ha.

In the second case from this region (example #3), the participating farmer was further inland with much lower yield expectation than in the previous example. The paddock planted to wheat in 2002 consisted of a shallow sandy loam soil with small water-holding capacity and low soil water content before crop establishment (Table 2).

Similar to the meeting process at Mingenew, pre-season yield predictions in response to crop management were provided to the farmer group at Dalwallinu. A wide range of yield expectations was simulated at this time. However, here we concentrate on the farmer meeting held on 12 August 2002 and the question then raised of whether the farmers should tactically invest further in crops that had established poorly. Management options included topdressing fertiliser and vigorously controlling weeds. The process for running the simulations was, as before, using actual soil and climate conditions for the 2002 season up until 12 August and completing the simulations with historical climate data for Dalwallinu thereafter.

APSIM simulations on 12 August 2002 strongly indicated that the 2002 season wheat crop sown on the nominated paddock had little chance of success: there was only a 15% chance of harvesting any grain and there was no prospect of yielding greater than 1 t/ha (Fig. 2). In contrast, the simulated long-term yield distribution for a hypothetical continuous wheat scenario with the same agronomic inputs (but not the starting soil or rainfall conditions as found in 2002) resulted in a simulated median yield of 1.8 t/ha. This dire prediction for 2002 was provided over 100 days before the expected crop harvest and was driven by the negligible soil stored water and low early-season rainfall. This result gave a strong indication to the collaborating farmer to not invest added inputs into his crop. By the end of the season, the crop had only received 75 mm in-season rainfall and it yielded only 0.2 t/ha. APSIM simulated a final yield of 0.01 t/ha.

---

**Fig. 1.** (a) Probability of exceedence for grain yield for crops planted on 22 May every year with the same starting soil and rainfall conditions as measured in 2002. Probabilities are shown for all years (solid line) and for those years with a consistently negative SOI phase in May–June, the then 2002 situation (dashed line); and (b) change in grain yield with 50 kg/ha urea applied on 25 July 2002 to crops planted on 22 May every year.
As a consequence of this series of interactions, the participating farmer groups in Western Australia initiated their own exploration of the credibility and relevance of APSIM to their production systems and adopted it as an exploratory tool. Such use continues today (Weeks et al. 2007).

Yield forecasting

The 2006 winter season on the Yorke Peninsula in South Australia commenced with good early rainfall (171 mm total from January to June 2006) such that most farmers in the region planted crops early which grew well through early winter. However, the last significant rain fell in the middle of July (34 mm total for the month) and the August–September period was the driest on record (14 mm total). A risk management strategy used by many local growers under such circumstance is to cut intended grain crops for hay. This strategy takes advantage of the good vegetative growth that crops have made, and offsets the risk associated with relying on good finishing rains to achieve a profitable grain yield. As grain fetches a higher price per tonne than hay, such decisions require a good estimate of the likely grain yield.

In example #4, the participating farmer had subscribed to the Yield Prophet® (www.yieldprophet.com.au) forecasting system whereby APSIM yield predictions are updated on demand throughout the season for nominated wheat and barley crops (Hunt et al. 2006; Hochman et al. 2009b, this issue). In this Yield Prophet® application, the planned wheat crop was set up in the nominated paddock before the start of the season by the farmer selecting the closest meteorological station (Ardrossan), the soil profile characterisation for the chosen paddock (characterisation data had been established from this paddock in the previous season), and the wheat variety to be planted (Table 2). Data were put into the webpage on the soil water and nitrate-N levels measured before the season (Fig. 3) and on the requested agronomic information such as sowing date and pre-plant and top-dress N fertiliser applications. Prior to each forecast query, the farmer put in his actual farm rainfall records up to the then current reporting day. In response to a requested report, the Yield Prophet® system returns a probability of exceedence distribution of wheat yield. This was simulated by APSIM based on the simulation of the wheat crop in that actual season up until the day of reporting, combined with 100 season finishes represented by the past 100 years of climate record for the site being simulated.

Based on some dry-matter cuts in mid September 2006, and a ‘gut feel’ estimate of likely grain yield given no rainfall for the remainder of the season, the participating farmer initially estimated he would get a greater financial return if he cut one or several of his wheat crops for hay. However, before he

![Fig. 2. Probability of exceedance APSIM simulations for continuous wheat across the historical climate data for Dalwallinu and for the 2002 wheat crop as simulated on 12 August 2002.](image-url)

![Fig. 3. For example #4, soil sampling on 30 April 2006 of a paddock at Ardrossan: (a) crop lower limit (wheat, ○), drained upper limit (●), and measured soil water (■) – plant-available water at sampling on 30 April 2006 was –7 mm to the sampled depth of 1 m – and (b) available nitrogen as NO3 (●) and NH4 (■) (there was a total of 105 kg/ha of N available to the sampled depth of 1 m).](image-url)
proceeded, he checked the forecast yield of his paddock using Yield Prophet®. The resultant simulation (Fig. 4) indicated that even with a worst-case season finish (i.e. if the season finish of 2006 were to equal the driest on record), the wheat in this paddock would still achieve a grain yield of 2.7 t/ha. With the price of wheat increasing as the full extent of the nation-wide drought became clear, and providing Yield Prophet® predictions were accurate, his calculations suggested he would be better off to not cut the paddock for hay, and instead harvest it for grain.

Despite initial scepticism of the grain yields predicted by APSIM in this very poor season finish, the farmer put his faith in Yield Prophet® and decided not to cut the paddock. A worst-case scenario did eventuate with the poorest season finish on record. The APSIM prediction of a lowest yield of 2.7 t/ha, estimated on 18 September 2006, was higher than the average yield for the harvested paddock (1.8 t/ha). This relatively large differential resulted both from a season finish outside the then-record of climate over 100 years and from a soil characterisation which required adjustment for this paddock. When the soil characterisation was amended, the simulated yield of the crop at the end of the season was 1.6 t/ha. This case study illustrates two difficulties in delivering Yield Prophet® across many paddocks: first, the consequences of crops being affected by exceptional circumstances or phenomena not addressed by APSIM; and second, the necessity to characterise well the soil resource.

Despite the lower than predicted yield, the farmer in this case estimated that his decision (not to cut the crop for hay) was of significant economic benefit and he commented:

‘It’s fantastic. Having access to this kind of science to check against your gut-feel is just invaluable. Given the thousands of hectares of wheat cut for hay across much of South Australia this season, it is clear models such as Yield Prophet® can save the industry many millions of dollars by making decisions such as hay versus grain more clear cut. There are many instances where farmers have cut crops for hay only to find the grain yields in paddocks left standing far exceeded their expectations.’

He remains an enthusiastic user of APSIM today.

The Yield Prophet® forecasting system for broadacre grain crops (wheat, barley, sorghum) is currently offered as a subscription system nationally around Australia (Hunt et al. 2006; Hochman et al. 2009b, this issue).

**Scenario exploration**

On 23 September 1997, the FARMSCAPE team was asked by representatives of a Queensland farmer association to make presentations to a specially convened meeting for the banking sector. This meeting was initiated in response to a widespread perception that banks were either withholding funds or forcing clients into ill-advised planting decisions given the wide perception of an El Niño-induced drought. At that time, the SOI was very low (~14.8 average for September 1997), the SOI phase for the July–August period was rapidly falling (Stone and Auliciems 1992) and there were warnings in the media about the worst El Niño event in 100 years. However, in the lead up to this summer season, the FARMSCAPE team had distributed information via our agribusiness alliances and the local media, which suggested that yield prospects are not greatly affected if fallows have reasonable stored water (>50% PAWC) or, in lower fallow water conditions, if planting date is delayed in order to store soil moisture. Our farmer collaborators were motivated to expose the banking sector to these results and invited ~20 representatives from regional and national banks, from local farmer organisations, and agribusiness firms. The meeting was held on a farm on the Darling Downs in south-eastern Queensland.

Widespread soil sampling of individual fallow strips on the Darling Downs during August and September 1997 indicated that, despite the dry winter, soil water levels in many fallows contained reasonable to good water storage. As would be expected, land that was long fallowed (generally 12 months or more since the harvest of the previous crop) held the most water, with many soil profiles being full at 100% PAWC and the majority being above 60%. Short fallows, where the crop was harvested early in the previous summer (e.g. spring sorghum or mungbean), generally had moderate levels of stored water (50–60%) due to the increased capture of rainfall in January and February 1997. In contrast, after later harvested crops such as cotton and late sorghum, soil water levels were below 50% PAWC.

The question ‘what would have happened if I had done one thing rather than another in past years?’ is important in deciding what to do ‘this’ season. Analysis of rainfall records can help by telling us something about the odds of success. As was the case in August and September 1997, the SOI was consistently negative (i.e. Phase 1) in 18 of the last 100 years. At the bankers’ meeting, we contrasted the results of those years with all other years in the record to determine the significance of the SOI information when ‘filtered’ through the APSIM model. In example #5, with output presented in Fig. 5, yields were predicted for Dalby cotton crops planted on 15 October at 3 starting soil water contents (Table 2). Each year of the simulation was initiated with the same nominated conditions. When examining the resulting graphs, meeting participants were asked not to compare the simulated yields with the actual yields achieved in any particular year, because the conditions assumed in the simulation were probably different from the actual conditions experienced by any crop in that year. Rather, the question posed was: what would the crop have

![Fig. 4. Probability of exceedence for the wheat crop at Ardrossan as of 18 September 2006.](image-url)
yields in that year if it had been planted under the assumed conditions? By reviewing the yield likelihood over 100 seasons, including seasons in which the SOI phase in September was Phase 1, one can assess the odds of succeeding with alternative management options.

For cotton planted on 15 October each year, the large year-to-year variation in simulated yield was evident for all levels of stored soil water. However, the effect of moving from 25% to 75% starting soil water increased median yield from 0.3 bales/ha to over 6 bales/ha. By contrast the effect of the SOI phase, comparing Phase 1 and all other years, at any starting soil water, was small. When starting with 25% PAWC, SOI Phase 1 years were, on average, lower yielding than other years, but this effect dissipated when starting with higher soil water storage values. The slight positive effect of SOI Phase 1 years at 75% PAWC was interpreted as due to the lower incidence of waterlogging events in those years.

In the case of sorghum in this same situation, APSIM analyses indicated that Darling Downs farmers could improve grain yield prospects by delaying sowing at least until mid November. In 97 of the last 100 years there were sowing opportunities for sorghum after that date; the other 3 years were not in the then current SOI phase. More detailed presentation of this analysis of sorghum sowing date is reported by Hochman et al. (1998).

As it transpired, sowing rains occurred in early October 1997. At that time, the most important predictor of likely yield for cotton and other crops was the water stored in the soil. The SOI phase added further to our predictive ability, but for a summer crop the influence of SOI phase on yield expectation is relatively small.

The meeting was highly successful as judged by recorded comments at the meeting and in subsequent actions by participants. The bank and agribusiness representatives appeared to correctly interpret the value of stored water relative to the climate forecasts, as evidenced in one comment that the 'poorest yields at 75% starting soil water are still better than average yields when starting with 25% soil water'. The bank representatives requested an information package based on the results presented at the meeting for distribution through their systems. Subsequent to the meeting, significant project-level follow-on activities were undertaken with both bank and agribusiness supply companies in the region to present soil monitoring and systems simulation to their staff and clients. Description of such follow-on activities is reported by Brennan et al. (2000) and Hochman et al. (2002).

Discussion
A model as a representation of scientific knowledge or as a tool for scientific inference is judged to be good if its processes and outputs correspond adequately with empirical measurements. However, evaluation of a model as a practical simulator of farm paddock performance demands somewhat different criteria. Here, adequate correspondence with reality is not sufficient. Scientifically rigorous demonstration of models’ capabilities for accurate simulation is not the key to changing the attitude of those farmers who consider models to be merely ‘toys for scientists’. In practice, ‘goodness’ of a model is judged according to experience of its performance in aiding ‘thinking about’ (a) a particular management situation and (b) what to do. The requirement for assisting decisions is explored further by other papers in this series (McCown et al. 2009, this issue). The reality that is meaningful to a farmer is the situation in question rather than the abstract ‘crop’ of the modeller. But because value in the practice situation cannot be achieved if simulations do not adequately represent the specific realities, the assessments of farmers and management consultants provide an important, albeit unconventional, form of model validation. This paper differs from most reports testing the performance of cropping systems models. Here, the measured yields were from commercial crops not experimental plots, the described applications were from real-life situations faced by farmers,
and the ultimate judges of quality of model performance were these farmers and their advisers.

A consequence of model-based research in this interpretive, or action, paradigm is that documenting and reporting the most important results, i.e. the value experienced by farmers, does not lend itself to conventional scientific criteria. Nevertheless, after 17 years of learning through such engagements, we conclude that the accuracy of simulations of actual paddock yields has generally been adequate to engender farmers’ confidence in the simulator sufficiently to ‘experiment’ with alternative actions. And more generally in the farming community, models are seen in a much more favourable light in Australia than 17 years ago.

Explaining on-farm simulator performance

The conventional approach to developing crop simulation models has been to build and test them using small-plot experimental data. These are conditions where deviation from the regression line due to errors of determination is minimised. Management can virtually assure the correctness of assumptions of non-limiting levels of targeted nutrients or water. Spatial homogeneity ensures low sampling errors in measurement of soil parameters, initial conditions, and yield. Model validation under such conditions generally accounts for 70–90%, but can be as low as 60%, of the observed variation in crop yield (e.g. Hammer and Muchow 1991; Robertson et al. 2002). The expected performance for simulated yields of farm paddocks, where conditions presumably deviate drastically from experimental plots, would be for poorer model performance. Yet the regressions reported here of simulated farm paddock yields compare favourably with performance reports under experimental conditions.

We submit two complementary explanations for this surprising result. Both concern departures in FARMSCAPE from norms regarding assumptions in simulations about on-farm production conditions. The first concerns our investment in specifying the simulator with information from the actual paddock in question and use of site-specific weather data rather than use of more available and less-costly surrogate information. Previous reports where paddock-specific data were used in simulations showed similar results to those reported here (Sadrás et al. 2003; Whish et al. 2005, 2007). Yet, investment in customised simulation is uncommon. In simulation of farm yields in Argentina by Mercau et al. (2007), the CROPGRO-Soybean model explained only 34% (original model) and 41% (calibrated model) of the measured variation in grain yields. There, soil characterisation and initial conditions were not determined in the field but rather estimated from soil survey data and expert opinion. The emphasis and investment in describing and parameterising the soil resource as a source of water and nutrient supply to crops have been largely ignored in crop modelling efforts to date. Crop physiology and crop modelling have been dominated by concern with plant growth and above-ground processes. The fact that Sinclair and Seligman (2000) ignored soil resource supply in attempting to establish criteria for reporting on crop modelling is an indicator of such misplaced emphasis. The biggest constraint to simulating crop yields in Australia remains access to paddock-specific soil information.

The second explanation for the surprising degree of agreement of simulations with farm yields can be related to the management performance of Australian farmers. Prior to the modelling era, discrepancy in yields between farm and experimental conditions was given prominence in Australia by Davidson (1962). He showed that mean weighted farm yields reached only 60% of experimental yields for wheat in Australia. More recently, Sadrás and Angus (2006) found that Australian wheat yields in experimental plots were 30% higher when compared with comparable large fields. This disparity is often referred to as the ‘yield gap’ between potential yields and farmer yields, and significant research effort has been aimed at closing this gap (Rockström and Falkenmark 2000; Dobermann and Cassman 2002; Huang et al. 2002). By inference, it is to be expected that yields simulated by computer models, built and tested on small-plot experimental data, will significantly over-predict paddock yields on commercial farms. In their review of yield potential in cereal crops, Evans and Fischer (1999) both inferred and implied as much by equating simulated yields to potential yields.

The results reported in this paper indicate that the commercial yields of Australian farmers, such as those sampled in this study, often approach ‘experimental’ or potential yields, i.e. there is a low yield gap. It appears that, on these farms, most crops are not seriously affected by yield-reducing factors such as weeds, pests and diseases, or poor agronomic management that may not allow crops to meet their potential. If such factors were influential, APSIM simulations, which neither explicitly account for biotic stresses nor for poor operational management, would significantly over-predict crop yields. Twenty years ago, Martin et al. (1988) found that most variation in wheat grain yield in Australia’s north-eastern wheatbelt could be related to available soil water and soil nitrate at sowing, sowing date and, significantly, weed density. As with weeds, cereal root diseases caused significant yield loss in past cropping systems, but this situation has been significantly improved through the introduction of break crops into cereal rotations and resistant varieties (Angus and van Herwaarden 2001). This study, using APSIM simulations of over 700 crops Australia-wide, suggested that the supply of water and nitrogen (variables well dealt with by APSIM) can account for most of the variation in crop yield, while farmers are mostly controlling other potentially yield-limiting factors such as P nutrition, weeds, and diseases. Additionally, the economic realities of the cost-price squeeze on the relatively unsubsidised Australian agriculture are likely incentives to ensure efficiencies in operational management (Kingwell and Pannell 2005).

The practicalities of paddock-specific simulation for farmers

Some have seen the limited availability of site-specific input data as precluding simulation models as a means of efficiently managing crop agronomy (Dobermann and Cassman 2002). Fortunately, climate data suitable for application within crop simulation are readily available online for many stations in Australia within the SILO patched point dataset (Jeffrey et al. 2001; www.bom.gov.au/silo). However, this same situation does not exist for the critical soil properties required for crop
modelling. Some progress has been made in creating an Australian national database of soil properties within the Australian Soil Resource Information System (ASRIS) (Johnston et al. 2003; www.asris.csiro.au), although critical parameters, such as field-determined PAWC, are not currently incorporated.

Field-determined PAWC is not a new concept to research agronomists around Australia. Soil physicists, in the past, have related PAWC to physical measures of water content determined in a laboratory (e.g. −33 and −1500 kPa matric potential) (e.g. Mullins et al. 1981). However, several researchers have pointed to errors in using an approach which does not account for differences in crop type or for soils that exhibit variable depth and degree of subsoil wetting (Dubbele et al. 1982; Gardener et al. 1984; Thorburn and Gardner 1989; Dalgliesh and Foale 1998). PAWC is now generally accepted as determined from field measurement of wet and dry soil profiles (Dalgliesh and Foale 1998; Dalgliesh et al. 2009). There have been several attempts at predicting PAWC for soils using readily available soil survey data. Ahem (1988) assessed the predictive abilities of three approaches, based on cation exchange capacity, matric potential, and soil depth, with some success. Littleboy (1997) developed a method for estimating PAWC indirectly from particle size analysis, but this approach failed to distinguish differences between crops. It is also of limited value because particle size to depth is expensive to measure. Hochman et al. (2001) developed close relationships between the PAWC of farmers’ soils and soil type, crop species, and soil depth, but these parameters required access to extensive soil water data and represented only two main soil classes: the grey and black Vertosols. Ladson et al. (2002) used the FARMSCAPE database and other data to compare PAWC estimates with the Atlas of Australian Soils and found significant differences. They concluded both that the field measure of PAWC is essential and that a collated PAWC database was lacking in Australia, in contrast to other soil properties.

The FARMSCAPE program has seen investment in input data as an essential requirement for model applications in real-world agriculture. Consequently, FARMSCAPE has promoted regular deep soil coring throughout Australia and, in some regions, has helped uncover previously unknown soil water and nitrogen resources at depth in the profile (Ridge et al. 1996; Dalgliesh et al. 2009). To assist in this promotion, Dalgliesh et al. (2009) report efforts to develop standard methodologies for soil characterisation and monitoring in the dryland cropping regions of Australia, and to assist farmers and agronomists in their application. An end result of such effort is the APSOIL national database of soil properties suited as input into simulation models such as APSIM. This APSOIL database is being integrated with the Australian Soil Resource Information System (ASRIS) (Johnston et al. 2003).

The cases of practical validation

A further avenue for establishing credibility for APSIM has been support of its use in the farm consulting industry. Reporting the results of comparisons between simulated and measured yields is an important requirement in such on-farm activity and has become a regular requirement for the Yield Prophet® system in Australia (Hunt et al. 2006). Most of the published reports of APSIM on-farm applications have provided, at a minimum, estimates of prediction accuracy and some benchmarking activity (Robertson et al. 2000; Whish et al. 2005, 2007). Working with consultants and marketers, Robertson et al. (2000) benchmarked mungbean production against APSIM simulations and identified machine harvesting as a significant source of yield loss with current management. They also confirmed that mungbean could be successfully planted in spring in south-eastern Queensland. Likewise, Whish et al. (2005) confirmed APSIM’s ability to simulate sorghum production in a range of row configurations before establishing management rules using the model to guide recommendations on when sorghum should be planted in skip rows. Whish et al. (2007) undertook a similar process in determining planting rules for chickpea in the drier regions of northern Australia. In each of these three cases, public and private agronomists participated in the research (and co-authored the journal papers), with lessons put into practice through actual agronomic recommendations to their clients.

Carberry et al. (2002) assert that for a model to be used in farm management decision-making it needs to be flexible and comprehensive in its capability to address relevant issues. The purpose of the five application examples described previously is to emphasise that credibility is gained in the context of specific circumstances. Dealing with immediate questions of involved farmers, contextualising the simulations to the specific situation in question, providing simulation outputs in an iterative process, and together reviewing the ensuing seasonal results against provided simulations represent the process for creating mutual understanding and credibility. Of course, providing simulation outputs that guide practice change which, in turn, proves beneficial is a tremendous credibility-forming exercise (as per example #4). However, disparity between simulation outputs and seasonal results can also provide an opportunity for building of credibility. In both examples #3 and #4, yield predictions throughout the season demonstrated reasonable yield expectations, yet the seasons ended as among the worst on record. When such disparities emerged, simulations were re-run and updated output provided to the participating farmers, with explanation for the discrepancies sought. When using a simulator with a good ‘track record’, a particular disparity between simulations and actual crop performance should be seen as opportunity for further investigation and insight. The willingness of farmers to accommodate such opportunity demonstrates their dual criteria for credibility in a model: the need for adequate correspondence with reality as well as experience with the model aiding decisions.

At the aforementioned meeting with representatives of the banking sector, an early and longstanding farmer collaborator within the FARMSCAPE program gave a wonderfully succinct description of the history of FARMSCAPE. He said, in part, that:

‘It started as researchers coming out asking ‘what is wrong with our model?’ Now it is us farmers asking the model ‘what is wrong with my crop?’’

Conclusions

This paper has presented evidence that APSIM can predict the performance of commercial crops at a similar level to that
reported for its performance against experimental yields. The essential requirement for simulating commercial yields across the Australian dryland cropping regions is to accurately describe the resources available to the crop being simulated, particularly the soil water and nitrogen resources. Thus, the FARMSCAPE program has well demonstrated that APSIM is relevant to commercial crops in Australia by testing simulated yield against commercial yield and by applying APSIM in the context of real-world practice.

Acknowledgments

We thank the many farmers and agribusiness advisers involved in FARMSCAPE and Yield Prophet® projects for their enthusiastic and vital participation. The financial support provided by GRDC and other funders across many research projects is greatly appreciated.

References


Davidson BR (1962) Crop yields in experiments and on farms. Nature 194, 458–459. doi: 10.1038/194458a0


Re-inventing model-based decision support with Australian dryland farmers


Mullins JA, Donnallan IE, Vandervee BE, Berndt RD (1981) The plant available water capacity of the important agricultural soils of the uplands of the eastern Darling Downs. Queensland, Department of Primary Industries, Brisbane, Division of Land Utilisation, Engineering Services Section; Report 81/3.


Manuscript received 8 February 2009, accepted 20 July 2009

http://www.publish.csiro.au/journals/cp