Evaluation of the APSIM model in cropping systems of Asia


a CSIRO Agriculture and Food, Australia
b CIMMYT-India, NASC Complex, DPS Marg, New Delhi 110012, India
c Bangladesh Agricultural Research Institute (BARI), Gazipur, Bangladesh
d Climate Change, Alternate Energy and Water Resources Institute (CAEWRI), National Agricultural Research Center (NARC) NH, Shalzad Town, Park Road, Islamabad, Pakistan
e Bangladesh Rice Research Institute (BRRRI), Gazipur, Bangladesh
f Natural Resources Division, Pakistan Agricultural Research Council, Islamabad, Pakistan
g ICAR Research Complex for NEH Region, Muzfread, Mohala 793 109, India
h Nepal Agricultural Research Council, National Grain Legumes Research Program, Rampur, Chitwan, Nepal
i Central Research Institute for Dryland Agriculture (CRIDA), Hyderabad, India
j Nanjing Agricultural University, Nanjing, Jiangsu 210095, China
k Field Crops Research and Development Institute, Department of Agriculture, Mahailulapallama, Sri Lanka
l International Rice Research Institute (IRRI), Los Banos, Philippines
m Agriculture Research and Development Centre, Bhub, Bhuban
n Rice Research and Development Institute, Department of Agriculture, Rakhulagoda, Libagamuwa, Sri Lanka
o ICAR – Indian Institute of Soil & Water Conservation, Dehradun, India
p ICAR – Indian Institute of Farming Systems Research, Muripuram-250110, Meerut, Uttar Pradesh, India
q Balai Pengajaran Teknologi Pertanian Nusa Tenggara Barat (BPTPNTB), Indonesia
r Faculty of Agriculture, University of Peradeniya, Peradeniya, Sri Lanka
s Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
t College of Resources and Environmental Sciences, China Agricultural University, Beijing, China
u Central Soil Salinity Research Institute, Karnal, Haryana, India
v CSIRO Land and Water, Brisbane, Australia
w School of Agriculture and Food Sciences, University of Queensland, Brisbane, Australia

A R T I C L E   I N F O

Article history:
Received 29 August 2016
Received in revised form
15 December 2016
Accepted 20 December 2016

Keywords:
APSIM
Asia
Cropping systems
Rice
Wheat

A B S T R A C T

Resource shortages, driven by climatic, institutional and social changes in many regions of Asia, combined with growing imperatives to increase food production whilst ensuring environmental sustainability, are driving research into modified agricultural practices. Well-tested cropping systems models that capture interactions between soil water and nutrient dynamics, crop growth, climate and farmer management can assist in the evaluation of such new agricultural practices. One such cropping systems model is the Agricultural Production Systems Simulator (APSIM). We evaluated APSIM’s ability to simulate the performance of cropping systems in Asia from several perspectives: crop phenology, production, water use, soil dynamics (water and organic carbon) and crop CO2 response, as well as its ability to simulate cropping sequences without reset of soil variables. The evaluation was conducted over a diverse range of environments (12 countries, numerous soils), crops and management practices throughout the region. APSIM’s performance was statistically
assessed against assembled replicated experimental datasets. Once properly parameterised, the model performed well in simulating the diversity of cropping systems to which it was applied with RMSEs generally less than observed experimental standard deviations (indicating robust model performance), and with particular strength in simulation of multi-crop sequences. Input parameter estimation challenges were encountered, and although ‘work-arounds’ were developed and described, in some cases these actually represent model deficiencies which need to be addressed. Desirable future improvements have been identified to better position APSIM as a useful tool for Asian cropping systems research into the future. These include aspects related to harsh environments (high temperatures, diffuse light conditions, salinity, and submergence), conservation agriculture, greenhouse gas emissions, as well as aspects more specific to Southern Asia and low input systems (such as deficiencies in soil micro-nutrients).

Crown Copyright © 2016 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The world will need 70–100% more food by 2050 (World Bank 2008; Royal Society of London, 2009), and current trends in population and consumption growth will mean the global demand for food will increase for at least another 35–40 years (Godfray et al., 2010). Projections indicate that the production of cereals must increase by roughly 2% per annum over the next four decades to ensure food security in South Asia (Ray et al., 2013). From a sectoral perspective, the production of rice, wheat, and maize must increase by about 1.1%, 1.7%, and 2.9% per annum, respectively (Gathala et al., 2014). To meet this demand sustainably, crop intensification while increasing resource-use efficiency and reducing the environmental footprint, or ‘ecological intensification’ (Cassman, 1999; Cassman et al., 2003; Ladha et al., 2009; Hochman et al., 2013) or ‘sustainable intensification’ (Royal Society of London, 2009) will be obligatory. For example, the Indian Punjab has been heralded for its technical achievements in past decades but increasingly criticized for leveraging its success on the environment (Jalota et al., 2007).

More recently, the debate about sustainable intensification is being extended to include aspects of resilience to climate change through the Climate Smart Agriculture framework (Campbell et al., 2014). Achieving such gains in productivity whilst reducing degradation of environmental resources will require a holistic systems approach, potentially incorporating the principles of conservation agriculture (CA), and judicious crop rotations (Hobbs 2007; Balasubramanian et al., 2012), amongst other potential adaptations. To complicate matters, any system advancements must be achieved under the overbearing shadow and uncertainty of a changing climate (Godfray et al., 2010). Cropping systems models have previously been used to explore the extent to which practices to manage climate risk can be judged to be climate smart (Hochman et al., 2017b).

There is a desire to investigate new practices in Asia with the aim of enhancing water productivity (WP) (Bouman, 2007), and cropping intensity (Dobermann and Witt, 2000) whilst maintaining environmental sustainability (Humphreys et al., 2010). Suggested pathways include the incorporation of non-flooded crops and pastures into traditional rice rotations (Cho et al., 2003; Singh et al., 2005; Zeng et al., 2007), changed agronomic and/or irrigation practices (Bouman and Tuong, 2001; Belder et al., 2007; Sudhir-Yadav et al., 2011a; Gathala et al., 2014), reduction of non-productive water losses (Humphreys et al., 2010), and genetic improvement (Peng et al., 1999; Sheehy et al., 2000; Bennett, 2003). Well-tested and locally-calibrated and validated simulation models are useful tools to explore opportunities within the context of a holistic systems approach – for increasing system productivity, assessing environmental trade-offs, and evaluating the effects of a changing climate.

Models such as DSSAT (Jones et al., 2003), ORYZA2000 (Bouman and van Laar, 2006) and Infocrop (Aggarwal et al., 2006a,b) have been widely used and tested in Asia, performing valuable studies on topics such as yield gaps and yield trends, evaluation of sowing dates and crop varieties, nitrogen and water management, environmental outcomes, and assessment of climate change impacts (Timsina and Humphreys, 2006a; Krishnan et al., 2007; Devkota et al., 2013). However for more in-depth evaluation of future farmer management strategies, there is also the need for a model which can more effectively simulate the performance of real farmers’ decision-trees and management which changes from year-to-year in response to unfolding climate and environmental conditions. None of the aforementioned models is able to do this.

The APSIM cropping systems framework (Keating et al., 2003; Holzworth et al., 2014) is such a model, with a proven track record in modelling the performance of diverse cropping systems, rotations, fallingow, crop and environmental dynamics (Turpin et al., 1998; Carberry et al., 2002; Robertson et al., 2001; Verburg and Bond, 2003; Whitbread et al., 2010; Hochman et al., 2014). A distinctive innovation and philosophical departure from most other ‘crop models’ is APSIM’s primary focus on simulating crop resource supply (rather than a primary focus on resource demand), with the soil forming the central simulation component. Crops, with their own resource demands impacted by weather and management, find the soil in one condition, and leave it in another condition for the next crop (McCown et al., 1996). This emphasis on simulation of soil resource dynamics positions APSIM strongly in comparison with other models for investigations into long-term changes to soil conditions and sustainability associated with different cropping strategies and practices. With particular focus on research into adaptation strategies, another notable strength of the APSIM model is it’s unique capacity to capture intricate detail and subtleties of dynamic farmer management practices through a highly flexible ‘Manager’ Module allowing the user to specify detailed farmer decision-trees in simple ‘if-then-else’ logic (Holzworth et al., 2014). APSIM has recently been enhanced to simulate rice-based cropping systems and environmental dynamics of ponded systems (Gaydon et al., 2012a,b). Evaluation of APSIM is well-established and well-documented in Australia and Africa, however limited in Asia as the model’s capability to simulate rice-based systems in this region is relatively recent. Also, rice is grown in more diverse cropping situations than most other crops (lowland (ponded), upland (un-ponded); puddled, un-puddled), necessitating more faceted evaluation. The first step in evaluating a model’s credentials is to define model capacities required for addressing research questions around some of the aforementioned issues. We suggest that a model for simulation of cropping system performance in Asia should be capable of several key functions: (i) robust crop development and yield simulation for a wide variety of crops; (ii) the ability to simulate cropping sequences and the effect of different fallow management, tillage and residue management strategies on system performance; (iii) robust simulation of soil water and nutrient dynamics in conjunction with crop performance; (iv) flexibility to capture detailed farmer-imposed management practices,
including subtle changes to farmer decisions and strategies, and evaluate their impact on system performance; and (v) robust simulation of crop response to CO₂ and temperature variation (Rötter et al., 2011). Before a researcher can be confident in employing such a model, it should be rigorously tested in a wide variety of environments and management practices.

Several large agricultural research initiatives in the Asian region of recent years have facilitated a broad assessment of the APSIM model, its strengths, weaknesses and priorities for future development. These include the ‘Adaptation to Climate Change in Asia’ project (ACCA; Roth and Grünbühel, 2012; Van Wensveen et al., 2016) and a project with the South Asian Association of Regional Cooperation to train south Asian scientist in the use of APSIM (SAARC-Australia Project; Gaydon et al., 2014b), the National Science and Technology Support Program of China, the National Basic Research Program of China, and the CSIRO-Chinese Ministry of Education (MOE) PhD Research Fellowship Program (Chen et al., 2010a,b; Liu et al., 2013; Wang et al., 2014). Data collected as part of these as well as other associated projects formed the basis of the evaluation presented here.

In this paper, we present details of APSIM evaluation across diverse cropping systems in Asia and critically evaluate its performance, identifying strengths and weaknesses using 43 experimental datasets from 12 countries, covering a broad spectrum of management practices, crop species/varieties, and environments, evaluating the hypothesis that APSIM is a robust model for use in Asian cropping systems research. As part of this process, we document input parameter estimation challenges and indicated needs for further improvement of the model.

2. Methods

We followed a robust three step process for each experimental dataset to ensure best possible model configuration, before compiling and conducting the final evaluation statistics on model performance. These three steps were:

- **Parameterisation**: This refers to the process of supplying the model with local input parameters which were directly measured or recorded (climate variables, soil physical and chemical characteristics, management impositions and inputs, etc.)
- **Calibration**: Other system parameters, unable to be directly measured or whose values possess greater uncertainty (such as crop varietal coefficients, some difficult-to-measure soil parameters etc.) required iterative adjustment, or calibration. In this process, the parameterised model was run for a chosen single year and treatment of the experiment using the best initial estimates for those input parameters to be calibrated. Simulated outputs (for crop development, production, soil water dynamics, or other variables of interest) were then compared with observed values from the experiment. If discrepancies were noted, parameters were re-adjusted within reasonable bounds, and the process repeated until acceptable model performance was achieved.
- **Validation**: The true test of the calibration was in testing the model against independent years and/or treatments of the experiment using the derived model calibration from step 2. This step is referred to as validation – essentially, a process to check the veracity of the model after the calibration and parameterisation of the previous steps. In complex models there is always the risk that an un-validated calibration can result in a model which is getting the right answers for the wrong reasons (Gaydon et al., 2012a), thereby leading to misleading model predictions during subsequent model usage. We performed validation simulations for all experiments. Once again this process became iterative – poor initial validation performance sometimes required revisitation of parameterisation and calibration steps, until acceptable model performance was achieved. Our definition of ‘acceptable’ is described in detail below (Section 2.3).

Evaluation statistics were performed on the compiled validation results to provide an insight into model performance and limitations across the breadth of our experimental datasets and their treatments. The following sections describe the process by which this model evaluation has been undertaken in detail.

2.1. Overview of the APSIM model

Detailed descriptions of APSIM are provided by Holsworth et al. (2014) and Keating et al. (2003). Here we merely provide a brief outline. APSIM is a dynamic daily time-step model that combines biophysical and management modules within a central engine to simulate cropping systems. The model is capable of simulating soil water, C, N and P dynamics and their interactions within crop/management systems, driven by daily climate data (solar radiation, maximum and minimum temperatures, rainfall). Daily potential production for a range of crop species is calculated using stage-related radiation-use efficiency (RUE) constrained by climate and available leaf area. The potential production is then limited to actual above-ground biomass production on a daily basis by soil water, nitrogen and (for some crop modules) phosphorus availability (Keating et al., 2003). The soil water balance (SOILWAT) module uses a multi-layer, cascading approach for the soil water balance following CERES (Jones and Kiniry, 1986), however a more process-based soil water–balance module is also available (SWIM3; Huth et al., 2012). The SURFACEOM module simulates the fate of the above-ground crop residues that can be removed from the system, incorporated into the soil or left to decompose on the soil surface. The SOILN module simulates the transformations of C and N in the soil. These include organic matter decomposition, N immobilization, urea hydrolysis, ammonification, nitrification and denitrification. The soil fresh organic matter (FOM) pool constitutes crop residues tilled into the soil together with roots from the previous crop. This pool can decompose to form the BIOM (microbial biomass), HUM (humus), and mineral N (NO₃ and NH₄) pools. The BIOM pool notionally represents the more labile soil microbial biomass and microbial products, whilst the more stable HUM pool represents the rest of the soil organic matter (SOM) (Probert et al., 1998). APSIM crop modules seek information regarding water and N availability directly from SOILWAT and SOILN modules, for limitation of crop growth on a daily basis. Biological and chemical processes occurring in ponded rice fields are simulated using the POND module within APSIM (APSIM-Pond, Gaydon et al., 2012b). Crop modules specifically relevant to the evaluation presented in this paper are APSIM-Oryza (Gaydon et al., 2012a), APSIM-Wheat (Wang et al., 2003), APSIM-Maize (Carberry and Abrecht, 1991), APSIM-Ozcot (Hearn, 1994); APSIM-Soybean (Robertson et al., 2001; Robertson and Carberry, 1998) and APSIM-Canola (used also for mustard; Robertson et al., 1999; Robertson and Lilley, 2016). APSIM-Oryza was recently improved to simulate rice crop response to soil salinity (Radaniello et al., 2016). APSIM-Wheat simulates salinity effect in a more simplistic way (focussing on water-availability effects only; Hochman et al., 2007) however none of the other APSIM crop modules attempt to simulate crop response to saline soil conditions.

2.2. Description of datasets

Datasets were assembled from across twelve countries in Asia; Bangladesh, Bhutan, Cambodia, China, India, Indonesia, Japan, Laos, Pakistan, Philippines, Nepal, and Sri Lanka. Essential criteria included the availability of detailed information on experimental
<table>
<thead>
<tr>
<th>Dataset No.</th>
<th>Reference</th>
<th>Location</th>
<th>Years</th>
<th>Crops</th>
<th>Estab</th>
<th>Treatments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Buresh et al. (1988)</td>
<td>Pila, Philippines</td>
<td>1985–86</td>
<td>Rice (IR58)</td>
<td>PTR</td>
<td>Five (5) N treatments (0.30,0.60,9.0,120 kgN/ha)</td>
</tr>
<tr>
<td>2</td>
<td>Suriadi et al. (2009)</td>
<td>Lombok, Indonesia</td>
<td>2007–09</td>
<td>Rice (ciguelis)</td>
<td>PTR</td>
<td>R-R-S-R-R rotation; two (2) water treatments (CS,AWD); three (3) N treatments: 0.69,138 kgN/ha/crop</td>
</tr>
<tr>
<td>3</td>
<td>Boling et al. (2004)</td>
<td>Java, Indonesia</td>
<td>1997–2000</td>
<td>Soybean (wilis) Rice (IR64)</td>
<td>DS</td>
<td>Six (6) consecutive rice crops (wet (DSR) and dry (PTR) seasons); three (3) N treatments: 0.120,144 kg N ha⁻¹ crop⁻¹ ; two (2) water treatments (LRF)</td>
</tr>
<tr>
<td>4</td>
<td>Bucher (2001)</td>
<td>Los Baños, Philippines</td>
<td>1996–2000</td>
<td>Rice (IR72)</td>
<td>PTR</td>
<td>Seven (7) consecutive rice crops; +- straw, early/late straw incorporation</td>
</tr>
<tr>
<td>5</td>
<td>Sudhir-Yadav et al. (2011a,b)</td>
<td>Punjab, India</td>
<td>2008–09</td>
<td>Rice (PAU-201)</td>
<td>PTR</td>
<td>Two (2) establishment methods (PTR, DSR); four (4) irrigation treatments (daily (CF), and AWD at soil water tensions (20cm) depth) of 20, 40 and 70 kPa. CF – one crop per year</td>
</tr>
<tr>
<td>7</td>
<td>Gaydon et al. (2013)</td>
<td>Gazipur, Bangladesh</td>
<td>2010–13</td>
<td>Rice (BRRI Dhan 28,29, 48, 49)</td>
<td>PTR</td>
<td>Maize (BARI Hybrid 7); Mustard (BARI Sarisha-15); Mungbean (BARI Mung-6)</td>
</tr>
<tr>
<td>8</td>
<td>Gaydon et al (unpubl)</td>
<td>Dacope, Bangladesh</td>
<td>2010–13</td>
<td>Rice (Nonakachi – local variety; BR23; BRRI Dhan 47); Cowpea (BARI Felon-1)</td>
<td>PTR</td>
<td>Four (4) cropping sequences; (i) fallow - fallow - T. Aman rice (Nonakachi); (ii) fallow - Fallow - T. Aman rice (modern variety-BR23); (iii) cowpea - fallow - T. Aman rice (BR23); and (iv) Boro rice (BRRI dhan47) - Fallow - T. Aman rice (BR23)</td>
</tr>
<tr>
<td>9</td>
<td>Gaydon et al (unpubl)</td>
<td>Satkhira, Bangladesh</td>
<td>2010–13</td>
<td>Rice (Nonakachi – local variety; BR23; BRRI Dhan 47); Cowpea (BARI Felon-1)</td>
<td>PTR</td>
<td>Four (4) cropping sequences; (i) fallow - fallow - T. Aman rice (Nonakachi); (ii) fallow - Fallow - T. Aman rice (modern variety-BR23); (iii) cowpea - fallow - T. Aman rice (BR23); and (iv) Boro rice (BRRI dhan47) - Fallow - T. Aman rice (BR23)</td>
</tr>
<tr>
<td>10</td>
<td>Gaydon et al. (2014a)</td>
<td>Satkhira, Bangladesh</td>
<td>2013–14</td>
<td>Rice (BRRI Dhan-47)</td>
<td>PTR</td>
<td>Three (3) salinity treatments: boro rice irrigated (CF) with (i) fresh water; (ii) mixed fresh and saline water; and (iii) saline water (dynamically changing in local canal)</td>
</tr>
<tr>
<td>11</td>
<td>Radanielson et al. (2016)</td>
<td>Infanta, Philippines</td>
<td>2013–14</td>
<td>Rice (IR64)</td>
<td>PTR</td>
<td>Four (4) salinity treatments; rice irrigated (CF) with (i) fresh water; (ii) alternate fresh and saline water (1 week intervals); (iii) alternate fresh and saline water (2 week intervals); and (iv) saline water</td>
</tr>
<tr>
<td>12</td>
<td>Radanielson et al. (2016)</td>
<td>Los Baños, Philippines</td>
<td>2012–13</td>
<td>Rice (IR64)</td>
<td>PTR</td>
<td>Four (4) salinity treatments; rice irrigated (CF) with (i) fresh water; (ii) 4 dS m⁻¹ ; (iii) 6 dS m⁻¹ ; and (iv) 12 dS m⁻¹ . Two crops per year (wet and dry seasons)</td>
</tr>
<tr>
<td>13</td>
<td>Rashid et al. (2009)</td>
<td>Chuadanga, Bangladesh</td>
<td>2005–07</td>
<td>Rice (BRRI Dhan-28, BR11)</td>
<td>DSR</td>
<td>Sowing date trial – two (2) seasons each of (i) T.Aman rice (BR11) and (ii) boro rice (BRRI Dhan-28). Five (5) sowing dates for T.Aman; Eight (8) sowing dates for boro.</td>
</tr>
<tr>
<td>14</td>
<td>Sena et al. (2014)</td>
<td>Karnal, India</td>
<td>2009–11</td>
<td>Rice (CSR36; CSR30)</td>
<td>PTR</td>
<td>Three (3) years; two (2) establishment treatments; two (2) varieties; three (3) N fertiliser rates (60, 80, 100 kgN ha⁻¹ (for CSR30); 120, 150, 180 kgN ha⁻¹ (for CSR30))</td>
</tr>
<tr>
<td>15</td>
<td>Choudhury et al. (2014)</td>
<td>Meghalaya, India</td>
<td>2009–11</td>
<td>Rice (Shahsarang)</td>
<td>PTR</td>
<td>Three (3) seasons; early transplant (7th July); 90 kg N ha⁻¹</td>
</tr>
<tr>
<td>Dataset No.</td>
<td>Reference</td>
<td>Location</td>
<td>Years</td>
<td>Crops</td>
<td>Estab</td>
<td>Treatments</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>16</td>
<td>Subash et al. (2014)</td>
<td>Modipuram, UP, India</td>
<td>2007–09</td>
<td>Rice (PR106)</td>
<td>PTR</td>
<td>R-W-R-W cropping sequence (4 crops)</td>
</tr>
<tr>
<td>17</td>
<td>Kumar et al. (2014)</td>
<td>Faizabad, UP, India</td>
<td>2000–03, 2010</td>
<td>Wheat (PBW343), Rice (Sarjoo-52)</td>
<td>DS</td>
<td>Three (3) dates of transplanting (10 June, 20 June, 30 June)</td>
</tr>
<tr>
<td>18</td>
<td>Rathnayake and Malaviarachchi (2014)</td>
<td>Batalagoda, Sri Lanka</td>
<td>2010–12</td>
<td>Rice (BG366, BG300, BG250)</td>
<td>DSR</td>
<td>Two (2) crops; three (3) varieties, rainfed conditions.</td>
</tr>
<tr>
<td>19</td>
<td>Suryagoda et al. (2006), Suryagoda and Pieris (2014)</td>
<td>Maha-Blup pallama, Sri Lanka</td>
<td>2006–07, 2011</td>
<td>Rice (BG300)</td>
<td>DSR</td>
<td>Irrigated with four (4) fertiliser applications (in kg urea ha⁻¹) – 12.5 (basal), 62.5 (14 das), 100 (42 das), 50 (56 das).</td>
</tr>
<tr>
<td>20</td>
<td>Amarasingha et al. (2015)</td>
<td>6 dry-zone locations, Sri Lanka</td>
<td>2000–11</td>
<td>Rice (BG300, BG359)</td>
<td>DSR</td>
<td>Recommended agronomic practice: varying dates of sowing; varying seasonal climatic conditions over 12 year period. Two sequential crops per year for seven years; several N treatment ranging from 0 to 100 kg N ha⁻¹ as well as a 10 t ha⁻¹ farmyard manure.;</td>
</tr>
<tr>
<td>21</td>
<td>Darai et al. (2014)</td>
<td>Nepalgunj, Mid-Western Terrai, Nepal</td>
<td>2003–07</td>
<td>Rice (Radha-4)</td>
<td>PTR</td>
<td>Four (4) rice varieties; recommended agronomic practice</td>
</tr>
<tr>
<td>22</td>
<td>Rai (2014)</td>
<td>Bhur, Bhutan</td>
<td>2010–11</td>
<td>Wheat (Gautam), Rice (RC88, IR72, IR780, Bhur Kambja2)</td>
<td>DS</td>
<td>Four (4) irrigation treatments – (i) irrigated to 100% of crop requirement; (ii) 80%; (iii) 60%; and (iv) non-irrigated/raifned</td>
</tr>
<tr>
<td>23</td>
<td>Ahmad et al. (2014)</td>
<td>Pothowar, Pakistan</td>
<td>2011–12</td>
<td>Wheat (Chakwal-50)</td>
<td>DS</td>
<td>Two (2) CO₂ treatments – (i) ambient (350–370 ppm) and (ii) FACE (580–645ppm) ; three (3) N treatments (4, 8, 12 g N m⁻²)</td>
</tr>
<tr>
<td>24</td>
<td>Kim et al. (2003)</td>
<td>Shizukuishi, Japan</td>
<td>1998–2000</td>
<td>Rice (Akitakomachi)</td>
<td>PTR</td>
<td>Two (2) CO₂ treatments – (i) ambient (370 ppm) and (ii) FACE (570ppm); three (3) N treatments (15, 25, 35 g N m⁻²)</td>
</tr>
<tr>
<td>26</td>
<td>Liu et al. (2013)</td>
<td>NCP, China</td>
<td>1981–2009</td>
<td>Rice (51 local cultivars)</td>
<td>PTR</td>
<td>Fertiliser application rate, number of irrigations, and +/+ residue</td>
</tr>
<tr>
<td>27</td>
<td>Wang et al. (2014)</td>
<td>NCP, China</td>
<td>1989–2003</td>
<td>Wheat (local vars) Maize (local vars)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>28</td>
<td>Chen et al. (2010a)</td>
<td>Luancheng, Hebei, China</td>
<td>1998–2001</td>
<td>Wheat (Gaoyou No. 503)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>29</td>
<td>Xiao and Tao (2014)</td>
<td>NCP, China</td>
<td>1998–2001</td>
<td>Wheat (cultivar-80, cultivar-00)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>30</td>
<td>Wang et al. (2013)</td>
<td>Tangyin, NCP, China</td>
<td>1990–1999</td>
<td>Wheat (local varieties)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>31</td>
<td>Chen et al. (2010b)</td>
<td>Luancheng, Yucheng, Fengqiu – NCP, China</td>
<td>1998–2006</td>
<td>Wheat (Gaoyou 503, Zhixuan 1, Keyu 13, Zhengmai 9023) Maize (Yandan 21, Yedan 22, 981, Zhengdan 958)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>32</td>
<td>Dong et al. (2014)</td>
<td>Northern China</td>
<td>1981–2003</td>
<td>Wheat (Gaoyou 503, Zhixuan 1, Keyu 13, Zhengmai 9023) Maize (Yandan 21, Yedan 22, 981, Zhengdan 958)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>33</td>
<td>Zhang et al. (2012)</td>
<td>NCP, China</td>
<td>2008–2010</td>
<td>Wheat (Nongda211, Han6172, Yanzhan4110)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>34</td>
<td>Wang et al. (2012)</td>
<td>Yangzte River Basin, China</td>
<td>2006–2008</td>
<td>Canola (Xiangzayou, Zhongshuang, Ningza)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>35</td>
<td>Liu et al. (2012)</td>
<td>North-East China</td>
<td>1983–2007</td>
<td>Maize (ten hybrids)</td>
<td>DS</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
<tr>
<td>36</td>
<td>Hochman et al. (2017a)</td>
<td>Andhra Pradesh, India</td>
<td>2011</td>
<td>Rice (WGL-14, BPT-5204, Kayya); Cotton (Neeraja, Ankur, Brahma)</td>
<td>PTR</td>
<td>Cultivar performance under unlimited water and N conditions over 29 years</td>
</tr>
</tbody>
</table>
soils, climate, imposed management, and observed crop phenology, final biomass and grain yield. Many datasets possessed additional measurements such as intra-crop biomass measurements, soil water and or nitrogen dynamics, and/or system water balance components (observed transpiration, evaporation, runoff, drainage etc.). Only datasets with replicated experimental data were used. Experiments covering at least two seasons were essential, for the purposes of both model calibration and subsequent validation. A broad spectrum of cropping environments and crop species/varieties across the region were represented (43 datasets; 966 crops — resulting in a model validation dataset composed of 361 rice crops, 326 wheat, 236 maize, and smaller numbers of cotton, soybean, mustard and canola crops; Table 1). The experiments encapsulated a diverse range of imposed treatments capturing a wide breadth of management practices in Asian agriculture. Several of these were multi-year, multi-crop sequences which included rice in rotations with other non-flooded crops, as well as wheat-maize rotations. Table 1 gives a description of each dataset used — geographical location, timeframe, treatments imposed and links to published references for further details.

2.2.1. Parameterisation and calibration protocol

APSIM requires daily values of rainfall, maximum and minimum temperature and solar radiation. Measurable soil physical parameters for different soil layers including bulk density, saturated water content, field capacity and wilting point are also required. Two parameters, U and CONA, which determine first and second stage soil evaporation (Ritchie, 1972) are also required. They were initially set at 6 mm and 3.5 mm day⁻¹, respectively, for the majority of datasets — values accepted for tropical conditions such as most of those described here (Probert et al., 1998; Keating et al., 2003). The proportion of water in excess of field capacity that drains to the next layer within a day was specified via a coefficient, SWCON, which varies depending on soil texture. Poorly drained clay soils will characteristically have values <0.5 whilst sandy soils that have high water conductivity can have values >0.8 (Probert et al., 1998). The values for saturated percolation rate (Ks in APSIM, mm day⁻¹) were extracted from published research papers add observation. Soil chemical parameters required by APSIM included soil pH, organic C and initial mineral N. The maximum daily algal growth rate for ponded conditions was estimated and assumed to be constant between sites (Gaydon et al., 2012b). Other parameters, not directly measured, required iterative calibration and are described in the following sections.

2.2.1.1. Soil organic matter (SOM) mineralization. Because SOM mineralization capacity varies between locations as a function of soil biota ecology and the proportion of SOM in the resistant pool (inert fraction), the values of the APSIM parameters Fbiom and Finert (Probert et al., 1998) were calibrated for each experiment using data from zero-N treatments, when available. In these treatments, a certain amount of plant-available mineral N was assumed to come from rainfall and/or irrigation water, and the remainder from mineralization of organic matter. In the absence of zero-N treatments, estimations were made using values from similar sites. The values of Fbiom and Finert were incrementally varied within physically plausible bounds (Probert et al., 1998) until the simulated indigenous N supply in the zero-N treatments allowed close simulation of the observed crop yields.

2.2.1.2. Soil carbon cycling. While in most cases only total soil organic carbon in the surface soil layer is available from measured data, APSIM requires initialisation of its SOM pools, i.e., fresh organic matter (FOM) pool, a more active carbon (BIOM) pool, and a humic (HUM) pool. The FOM pool contains all the fresh organic matter, such as dead crop roots and incorporated residues. The assumptions used to fractionate the total soil C in different soil layers and each soil C pool were similar to those of Luo et al. (2011): 1) total SOC, mostly concentrated in the top 20–30 cm, decreases exponentially with soil depth; 2) in deeper soil layers, the proportion of total soil C that is resistant to decomposition is higher, and the absolute amount of inert C is constant across the soil layers; and 3) the FOM pool is mainly distributed in upper soil layers and represents a small proportion of total SOC in the cropland soil with a long history of cultivation.

2.2.1.3. Crop phenology. In simulation of each experiment, crop varieties were calibrated by varying the APSIM crop phenology parameters until the modelled phenology dates matched the observed dates. Usually, the first crop in each dataset was used for this calibration procedure, and subsequent crops were used for model validation. The primary dates of focus were those associated with sowing, transplanting, maximum tillering, panicle initiation, flowering, and physiological maturity.
2.2.1.3. Crop biomass partitioning. Observed ratios between grain, straw, and straw components (stems, leaves etc.) were used to calibrate parameters governing allocation of assimilated biomass amongst plant components.

2.2.1.4. Soil water dynamics, ET and crop water uptake. Observed soil water dynamics and crop water uptake were used to calibrate crop root parameters (kfl and af in APSIM) and crop lower limits (cll) as a function of soil layer.

2.2.2. Validation protocols

The parameterised and calibrated model was then used to simulate independent years/season/crops in each experimental dataset, as a means of checking the veracity of the calibrations. These validation data-pairs (simulated vs observed) were used to evaluate the model's performance from a range of perspectives discussed in the following section.

2.3. Aspects evaluated

The model's capacity to simulate crop production (grain yields and above-ground biomass) for different crops (rice, wheat, maize, others), in different environments and under different management practices was the primary aspect evaluated. The combined dataset for rice crops was broken into subsets of rice establishment method (PTR- puddled transplanted rice; or DSR — direct-seeded rice (including both dry- and wet-seeded). This was overlayed with water management treatments comprising (i) irrigated with continuous flooding (CF); (ii) irrigated according to alternate wetting and drying (AWD); or rainfed (RF) lowland. The large size of the dataset allowed this (361 paired data points), and provided an opportunity for model evaluation for these key rice management options. At this stage in APSIM's application history in Asia, this breakdown process has not been possible for other crops due to restricted dataset size.

Other aspects of model performance evaluated were:

- The ability to robustly simulate crop phenology for sowing date trials.
- Simulation of crop sequences — by examining residual error as a function of crop progression (continuous simulation of multiple crops and fallows without resetting soil water and nitrogen variables between crops).
- Soil water and soil carbon dynamics — in conjunction with crop production (related to a model's ability to simulate cropping system 'carry-over' effects)
- System water balance components (crop transpiration, soil and pond evaporation, runoff, drainage etc.) and irrigation water use.
- CO₂ response — using free-atmosphere carbon enrichment (FACE) experiments

2.4. Statistical evaluation methods used

Linear regression was used to compare paired data-points for observed and simulated grain yield, for both rice and other crops. We determined the slope (α), intercept (β), and coefficient of determination (R²) of the linear regression between simulated and observed values. We also evaluated model performance using the Student’s t-test of means assuming unequal variance P(t), and the absolute square root of the mean squared error, RMSE (Eq. (1)).

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (O_i - \bar{S}_i)^2}{n}}
\]

Where S_i and O_i are simulated and observed values, respectively, and n is the number of data pairs. A model reproduces experimental data best when \( \alpha = 1 \), \( \beta = 0 \), \( R^2 = 1 \), and the absolute RMSE between simulated and observed values is similar to (and ideally less than) the standard deviation of experimental measurements (representing the error between treatment replicates, or the 'uncertainty' of the observed data). Statistical comparisons were conducted for subsets of the overall rice dataset, to explore the performance of the model in simulating different establishment practices (PTR and DSR) as well as different water management practices (irrigated, both fully ponded and AWD) and rainfed.

We also calculated the modelling efficiency, \( EF(Withmott, 1981; Krause et al., 2005) \) as another recognized measure of fit. The modelling efficiency is defined as:

\[
EF = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O}_i)^2}
\]

Where \( \bar{O} \) is the mean of the observed values. A value of \( EF = 1 \) indicates a perfect model (Mean squared error, MSE = 0) and a value of 0 indicates a model for which MSE is equal to the original variability in the observed data. Negative values suggest that the average of the observed values is a better predictor than the model in all cases.

3. Results

3.1. Crop production

A large proportion of our simulated datasets (23 out of 43) were simulated as cropping sequences of rice and non-rice crops over at least 2 seasons, with a further 10 representing sequences of rice-on-rice cropping, including fallows. Hence in the following section, even though the results are distilled and presented on the basis of individual crops, it should be noted that they implicitly represent the performance of the APSIM model in simulating crop sequences and system processes in which these crops were grown.

3.1.1. Rice

The major crop represented in our assembled datasets was rice, due to its dominance in Asian cropping systems. A large number of paired data-points (361 crops) allowed segregation into a range of management, and subsequent evaluation of APSIM model performance in simulating each of those across a range of geographical locations and environments. Considering the combined rice dataset as a whole, the model performed well in simulating above-ground biomass and grain yield (Fig. 1 and Table 2).

The RMSE of 1084 kg ha⁻¹ for the combined rice grain yield dataset compares favourably with the standard deviation amongst the observed data and replicates (2038 kg ha⁻¹). This is supported by a strong correlation between simulated and observed data \( r^2 = 0.83 \) with low bias \( \alpha = 1.1, \beta = -246 \) kg ha⁻¹. The Student's paired t-test (assuming non-equal variances) gave a significance of \( P(t) = 0.09 \), indicating that there is no statistical difference between observed and simulated data at the 95% confidence level, while the high overall modelling efficiency, \( EF \), of 0.72 indicates the model is performing acceptably. Hence there is convincing evidence here that APSIM is simulating rice yields well within the bounds of experimental error across a range of varieties, environments and imposed management practices, and its performance must be considered adequate over the range of this diverse Asian dataset.
Table 2
Statistics for observed vs simulated RICE grain yield (across different cultivation/irrigation practices)

<table>
<thead>
<tr>
<th>Est</th>
<th>Water</th>
<th>n</th>
<th>Xob (SD) (kg ha⁻¹)</th>
<th>Xsim (SD) (kg ha⁻¹)</th>
<th>P(t*)</th>
<th>α</th>
<th>β (kg ha⁻¹)</th>
<th>R²</th>
<th>RMSE (kg ha⁻¹)</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTR</td>
<td>CS</td>
<td>218</td>
<td>6250 (2065)</td>
<td>6619 (2563)</td>
<td>0.1</td>
<td>1.1</td>
<td>−250</td>
<td>0.83</td>
<td>1260</td>
<td>0.83</td>
</tr>
<tr>
<td>DSR</td>
<td>AWD</td>
<td>18</td>
<td>5385 (1222)</td>
<td>5047 (1513)</td>
<td>0.23</td>
<td>1.1</td>
<td>16</td>
<td>0.79</td>
<td>884</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>29</td>
<td>3547 (1417)</td>
<td>3777 (1867)</td>
<td>0.7</td>
<td>1.24</td>
<td>1037</td>
<td>0.88</td>
<td>714</td>
<td>0.88</td>
</tr>
<tr>
<td>(all PTR)</td>
<td>265</td>
<td>5896 (2131)</td>
<td>6218 (2634)</td>
<td>0.69</td>
<td>1.12</td>
<td>−379</td>
<td>0.83</td>
<td>1019</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>AWD</td>
<td>47</td>
<td>5232 (1103)</td>
<td>5308 (1112)</td>
<td>0.74</td>
<td>0.93</td>
<td>433</td>
<td>0.85</td>
<td>432</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>49</td>
<td>4254 (1590)</td>
<td>4555 (1789)</td>
<td>0.38</td>
<td>0.99</td>
<td>358</td>
<td>0.77</td>
<td>903</td>
<td>0.67</td>
</tr>
<tr>
<td>(all DSR)</td>
<td>96</td>
<td>4733 (1452)</td>
<td>4923 (1536)</td>
<td>0.38</td>
<td>0.95</td>
<td>444</td>
<td>0.79</td>
<td>712</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Overall Combined</td>
<td>361</td>
<td>5587 (2038)</td>
<td>5874 (2458)</td>
<td>0.09</td>
<td>1.1</td>
<td>−246</td>
<td>0.83</td>
<td>1084</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>

Est: crop establishment method (PTR – puddled transplanted rice; DSR – direct-seeded rice (incl. dry- and wet-seeded)); Water, water supply method (CS – irrigated and continuously submerged; AWD – irrigated with alternate wetting and drying; RF – rainfed lowland); Xob, mean of observed values; Xsim, mean of simulated values; SD, standard deviation; n, number of data pairs; P(t*), significance of Student’s paired t-test assuming non-equal variances; α, slope of linear regression between simulated and observed values; β, y-intercept of linear regression between simulated and observed values; R², square of linear correlation coefficient between simulated and observed values; RMSE, absolute root mean squared error; EF, the modelling efficiency.

* values greater than 0.05 indicates simulated and observed values are the same at 95% confidence level.

The analysis of the rice management subsets (Table 2) reveals no particular aspect in which APSIM has performed inadequately (all low RMSE with high modelling efficiencies), however the performance of the model in simulation of puddled transplanted rice (PTR) is better than direct-seeded rice (DSR) (P(t) of 0.69 cf 0.38, combined with EF of 0.85 cf 0.76). Across the water treatments, the simulation of continuously-ponded irrigation performs the best — not surprisingly due to APSIM-Oryza’s derivation from ORYZA2000 and CERES-Rice (ponded bio-chemistry and saturated soil environment algorithms) both of which were originally derived for PTR lowland irrigated rice production. This analysis indicates future model improvement efforts are best targeted at improved process-based simulation of PTR, alternate wet-and-dry (AWD) and DSR rain-fed (RF) (and probably DSR AWD, insufficient data available) systems.

3.1.2. Wheat, maize and other crops

In simulating regional experimental wheat yields, APSIM achieved a RMSE of 845 kg ha⁻¹ in a dataset with experimental standard deviation of 1794 kg ha⁻¹ (n = 326; Table 3, Fig. 2), once again comparable with CERES-Wheat which reported an RMSE of 480 kg ha⁻¹ for a dataset (n = 137) with SD of 1400 kg ha⁻¹ (for South Asia only). In conclusion, it appears the models are remarkably similar in their ability to provide acceptable simulation of individual rice and wheat crop performance in the region.
Table 3
Statistics for observed vs simulated WHEAT, MAIZE, COTTON, SOYBEAN, and MUSTARD grain yield.

<table>
<thead>
<tr>
<th>Crop</th>
<th>n</th>
<th>Xobs(SD)/(kg ha⁻¹)</th>
<th>Xsim(SD)/(kg ha⁻¹)</th>
<th>P(r)</th>
<th>α</th>
<th>β (kg ha⁻¹)</th>
<th>R²</th>
<th>RMSE (kg h⁻¹)</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHEAT</td>
<td>326</td>
<td>4397 (1794)</td>
<td>4272 (1818)</td>
<td>0.38</td>
<td>0.90</td>
<td>296</td>
<td>0.79</td>
<td>845</td>
<td>0.78</td>
</tr>
<tr>
<td>MAIZE</td>
<td>236</td>
<td>5972 (2408)</td>
<td>5973 (2523)</td>
<td>0.99</td>
<td>0.96</td>
<td>232</td>
<td>0.85</td>
<td>1004</td>
<td>0.83</td>
</tr>
<tr>
<td>COTTON</td>
<td>8</td>
<td>1769 (617)</td>
<td>1627 (603)</td>
<td>0.65</td>
<td>0.87</td>
<td>89</td>
<td>0.79</td>
<td>303</td>
<td>0.72</td>
</tr>
<tr>
<td>MUSTARD</td>
<td>10</td>
<td>1041 (485)</td>
<td>1433 (378)</td>
<td>0.06</td>
<td>0.57</td>
<td>836</td>
<td>0.55</td>
<td>502</td>
<td>−0.19</td>
</tr>
<tr>
<td>SOYBEAN</td>
<td>6</td>
<td>2152 (139)</td>
<td>2055 (84)</td>
<td>0.18</td>
<td>−0.44</td>
<td>2999</td>
<td>0.53</td>
<td>214</td>
<td>−1.83</td>
</tr>
<tr>
<td>CANOLA</td>
<td>19</td>
<td>2191 (769)</td>
<td>2026 (622)</td>
<td>0.47</td>
<td>0.68</td>
<td>545</td>
<td>0.71</td>
<td>444</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Xobs, mean of observed values; Xsim, mean of simulated values; SD, standard deviation; n, number of data pairs; P(r), significance of Student’s paired t-test assuming non-equal variances; α, slope of linear regression between simulated and observed values; β, y-intercept of linear regression between simulated and observed values; R², square of linear correlation coefficient between simulated and observed values; RMSE, absolute root mean squared error; EF, the modelling efficiency.

Maize, cotton and canola production were similarly simulated well, with RMSE values well within the range of observed experimental variability, P(r) indicating no significant difference between observed and simulated yields at the 95% level, and high modelling efficiencies (EF > 0.7 in all cases) (Figs. 2–4; Table 3).

Mustard and soybean performance was less robust (Fig. 4; Table 3) with lower confidence levels (P(r) and RMSE) figures very close to, or greater than, the observed experimental standard deviations. EF values were also negative, suggesting that the average of the observed values is a better predictor than the model in all cases. Given the strong validation of APSIM in simulating these crop species in more data-rich environments outside Asia (Robertson and Carberry, 1998; Robertson et al., 1999) these low figures are almost certainly due to the limited amount of validation data available (number of paired data points for Mustard (10) and Soybean (6)) combined with uncertainties on data quality). More validation work is indicated for these crops in Asia.

3.2. Crop phenology

Simulation of rice crop phenology for the two sowing date trials (datasets #13 and #41, Bangladesh and India respectively) indicated strong performance of APSIM in simulation of phenology (and associated yield responses) for both photoperiod-sensitive (Fig. 5) and non-photoperiod sensitive (Fig. 8) rice varieties. However the importance of good calibration for photoperiod sensitivity parameters (Nissanka et al., 2015) was clearly indicated (Figs. 6 and 7).

It is important to note that crops in these sowing date trials were fully irrigated, and neither water or nutrient-stressed. Water-stressed conditions (AWD and RF) are handled acceptably (high R², low RMSE; data not shown), however, deficiencies were noted in the simulation of N-stressed crops with significant bias indicating poor simulation of rice crop phenological responses to N stress (Table 4a and Table 4b).

Simulation of wheat and maize phenology for validation datasets indicated good correlation (r² > 0.99) and low error (Fig. 9), as did simulated results for other crops (data not shown).

3.3. Ability to simulate crop sequences (without resets)

The APSIM model was originally conceptualised and developed to simulate cropping sequences and allow research into the impacts on crop production of different fallow management practices and climate variability (Keating et al., 2003; Holzworth et al., 2014). To
Fig. 5. Simulated and observed data for sowing date trial of T.Aman rice, Chuadanga, Bangladesh (Dataset #13; Rashid et al., 2009). Graphs show a.) crop phenology; and b.) crop production, for a strongly photoperiod-sensitive rice cultivar (BR11). Simulated data represented by lines; observed data by points with error-bars representing the variability across replicates (one standard deviation either side of the mean).

Fig. 6. Sensitivity of APSIM-Oryza varietal photoperiod sensitivity parameter (PSSE) in simulation of rice sowing date trial, Chuadanga, Bangladesh (Dataset #13; Rashid et al., 2009); a.) PSSE = 0.2, b.) PSSE = 0.3, and c.) PSSE = 0.4.

Fig. 7. Sensitivity of APSIM-Oryza varietal optimum photoperiod parameter (MOPP) in simulation of rice sowing date trial, Chuadanga, Bangladesh (Dataset #13; Rashid et al., 2009); a.) MOPP = 11.0 h, b.) MOPP = 11.2 h, and c.) MOPP = 11.4 h.

Fig. 8. Simulated and observed data for a rice sowing date trial over two years, Andhra Pradesh, India (Rao and Reddy, 1998; dataset #41). Graphs show a.) crop phenology; and b.) crop production, for a non-photoperiod-sensitive rice cultivar (BPT-5024), in contrast to Fig. 5 (photoperiod sensitive).
the best of our knowledge, however, there has been no specific metric used in the scientific literature (relating to crop systems modelling) which allows evaluation of a model's ability to simulate system processes leading to successful modelling of cropping sequences. For example, Carberry et al. (1996) successfully simulated a multi-species, conservation agriculture farming system in Northern Australia, including impacts of mulch and legume-cereal sequences, but did not seek to evaluate model bias with time. We have chosen to evaluate the variation in residual error between simulated and observed grain yields as a function of progressive crop number in the sequence, to serve this purpose. In other words, the assumption is that residual error will not increase with successive crops if system processes are being correctly or adequately simulated. As cropping sequence datasets with appropriate data and the required degree of experimental rigour for all relevant variables are rare in Asia, we focussed on evaluation of three (3) most complete modelled datasets from our assembled selection (datasets #2 (Suriadi), #4 (Boucher), and #7 (Gazipur)), Fig. 10 gives a graphical example of simulated vs observed crop production from one treatment in each of these experiments.

All the simulated versus observed grain yield measurements for treatments across each of these three experiments were combined in an analysis of residual absolute error and the results illustrated in Fig. 11 as a function of advancing crop number in the sequence.

The data were used to estimate a linear regression of absolute error by the number of successive crops for each site. No coefficient estimates were both positive and significant indicating there is no evidence APSIM error increased with the number of successive crops modelled. By contrast, at sites ‘Suriadi’ and ‘Gazipur’ there is some evidence that the simulation error reduced (Fig. 11b; P(t)=0.096 and P(t)=0.054 respectively). This reduction in error may be due to the relative inadequacy of initial parameter estimates, which become less and less significant with each passing crop as system equilibrium is established and initial model conditions become less significant on crop performance. According to our assumption, this indicates robust simulation of system processes leading to confidence in APSIM performance in these rice-based systems.

3.4. Soil water dynamics and system water balance components

Experimental datasets in Asia with observed soil water dynamics and measurements of system water balance components (evaporation, runoff, drainage) in conjunction with crop performance are also rare, however growing in frequency with increasing accessibility and affordability of instruments and data-logging equipment. We used the two most complete datasets from the Indian Punjab to evaluate APSIM performance in simulating soil water dynamics under CF and AWD irrigated rice (dataset #5, Sudhir-Yadav, 2008-09) and under irrigated and rainfed wheat (dataset #38; Balwinder-Singh, 2006-08), as well as two datasets from China (dataset #28; Chen et al. (2010a), and dataset#31; Chen et al. (2010b)) under irrigated and rainfed wheat and maize.

The dynamics of soil water (Fig. 12a-c), water-balance components (Fig. 13a.) and associated crop production (grain yields, Figs. 14 and 15) were all simulated well, across a range of cropping systems and environments featuring rice, wheat and maize, and with contrasting stubble treatments (with and without). Accu-
racy of simulated biomass production generally decreased with increasing water stress level (Fig. 14), with over-prediction of crop performance indicating under-estimation of crop stress associated with AWD rice by the model.

3.5. Soil carbon dynamics

Confidence in APSIM’s ability to represent the sustainability of the diverse cropping practices encountered in Asia is indicated by simulation of sensible soil carbon dynamics in several experiments of more than 20 years duration, in diverse environments and cropping practices (dataset #27, Fig. 16, wheat-maize in China with a range of fertiliser and stubble treatments; and continuous flooded rice in the IRRI long-term cropping experiment, Gaydon et al., 2012b). There are several additional published studies from Asia demonstrating robust APSIM simulation of soil water, soil organic carbon dynamics and N mineralisation in association with crop yields for wheat and soybean cropping systems under different fertiliser and manuring practices (Bhopal, India; Mohanty et al., 2011, 2012).
3.6. Crop response to atmospheric CO2

Simulated crop response to increased CO2 levels is clearly an important criterion in model evaluation for future research needs in Asia relating to climate change. Only a small number of FACE experiments have been conducted with datasets amenable to our analysis. These were rice FACE datasets from Japan and China (datasets #24 and 25). Both rice biomass and grain yield predictions under ambient and increased atmospheric CO2 concentrations compared well with the experimental data (Fig. 17), and although no replicate variability data were available, we expect (based on standard error reported in similar trials from the research organisations concerned) that the model has simulated the observed crop responses within the bounds of experimental error.

There is no published validation of APSIM’s performance to simulate non-rice crop CO2 response in the Asian context, however there is evidence of adequate performance in simulating wheat
for FACE datasets in Australia (O’Leary et al., 2015) and the United States (Asseng et al., 2004).

3.7. Input parameter estimation challenges and ‘work-arounds’

The reporting of APSIM performance in this study is incomplete without noting a range of model parameter estimation challenges which were encountered and overcome — in many cases indicating model deficiencies and suggesting improvements required in model science. The issues and ‘work-arounds’ used are noted below:

3.7.1. N-stress effect on rice phenology

In dataset #1, there was an 11 day (6.4%) error in “time to maturity” under extreme N-stress conditions, a 4-day error under mid-range N-stress conditions, compared with simulation of an unstressed crop. This translated to a 12% error in total biomass production (extreme to no-stress). We worked around this problem by calibrating a new simulated ‘variety’ for each N-stress treatment with modified phenology parameters — calibrating, rather than modelling, the crop response to N-stress.

3.7.2. High temperature effect on wheat crop phenology

The default APSIM-wheat thermal time units increase linearly with increasing daily average temperature between 0 and 26 °C, reaching a maximum value at 26 °C which then decreases linearly to 0 again at 34 °C. We found this approach inadequate to adequately simulate the phenology of late-planted wheat crops in north-west India whose growth period experiences temperatures above the daily average temperature optimum of 26 °C. Using default parameters, APSIM overestimated the length of the crop growth period for such crops. After reviewing the literature and discussions with wheat physiologists, we modified this approach by plateauing the thermal time accumulation at a maximum for temperatures between 26 and 34 °C, declining thereafter to zero at 40 °C, similar to the approach now used in DSSAT (Jones et al., 2003, Ken Boote, personal communication). We tested the approach against available data and the modified model performed sensibly in simu-
lating wheat crop duration under such high temperature conditions (Balwinder-Singh et al., 2015a, 2016).

Chen et al. (2010b) also modified the temperature function for both thermal time and radiation-use efficiency (RUE). In their case, it was found that the original APSIM-Wheat overestimated the early growth after winter. So they modified the temperature response functions into curvilinear forms with reduced thermal time below 15°C and reduced RUE below 22°C based on Wang and Engel (1998).

3.7.3. High temperature effect on rice floret sterility

The APSIM-Orzya model (as per ORYZA2000 model) uses ambient temperature to calculate the fraction of infertile spikelets caused by high temperature. The average daily maximum temperature over the flowering period (0.96 ≤ crop stage variable (DVS) ≤ 1.22) is calculated. Average maximum temperatures above 36.6°C cause proportional floret sterility (Bouman and van Laar, 2006). The threshold temperature (htmax = 36.6°C) is a default APSIM-Orzya input parameter. In reality though the developing floret experiences the canopy temperature of the crop (not the ambient temperature), which has been demonstrated in dry environments to be as much as 6–7°C below ambient (fully-irrigated Australian rice at 20% relative humidity; Matsui et al., 2007). We used the relationship between ambient temperature and rice canopy temperature developed by Van Oort et al. (2014) (equation 10b in that paper) to modify the input value of htmax. This relationship is a function of ambient relative humidity (RH), so we estimated average RH values at each experimental location during the rice flowering phase, calculated the degree of canopy cooling from Van Oort et al. (2014) and increased the value of htmax accordingly. For example, for RH’s around or above 80%, the canopy cooling is zero hence htmax remained at the default value of 36.6°C. This covered most tropical datasets. An RH of 50% however results in a canopy cooling of around 3°C, hence in this situation we used htmax = 39.6°C (in other words, a higher threshold) to correctly simulate high-temperature floret sterility effects. Not employing this work-around resulted in overestimation of rice crop sterility in drier regions.

3.7.4. Low temperature effect on rice floret sterility

The effect of low temperatures on rice spikelet sterility (important for later sowing of longer duration varieties), is calculated...
according to Horie et al. (1992) based on the ‘cooling degree days’ concept, using a threshold temperature 22 °C. For north Indian simulations in this model evaluation, the default average daily temperature threshold in the model was changed from 22 °C to 28 °C based on the results of the field study of Nagarajan et al. (2010) conducted in a similar environment with similar rice varieties.

3.7.5. Mulch impacts on crop phenology

Balwinder-Singh et al. (2011) and Sidhu et al. (2007) demonstrated that wheat crops under mulched crop residues exhibit delayed anthesis by 6–8 days, but that APSIM crops do not simulate this phenomenon. The phenological parameters for wheat were therefore set independently for non-mulched and mulched treatments in our simulated datasets, so that anthesis was delayed by about 6 days in the mulched treatments, compared with the non-mulched crops.

3.7.6. Calibration of photoperiod response in rice crops

APSIM-Oryza demonstrated reliable simulation of rice crop phenology provided care was taken in parameterising and calibrating photoperiod sensitivity. Data from sowing date trials made this a relatively straightforward task, however the biggest issue for APSIM users in the region is adequately parameterising photoperiod-sensitive varieties when such data were not available (see Figs. 6 and 7, and Nissanka et al., 2015). In our experience from the ACCA and SAARC-Australia projects, the optimum photoperiod parameter (MOPP, hrs) is readily known and available for different rice crop varieties – the greater challenge is in selecting a sensible value for photoperiod sensitivity (PSSE). Our response was to use PSSE = 0.1–0.2 for a mildly photoperiod sensitive variety, 0.3 for a medium, and 0.4-0.5 for what is locally described as a strongly photoperiod-sensitive rice variety. If the rice crops were non-photoperiod-sensitive, the matter was simple (PSSE = 0, and MOPP was unused).

3.7.7. Submergence of young rice crops

When excess rainfall occurs during the early stages of a rice crop, resulting submergence can damage crops and delay phenology (Jackson and Ram, 2003). Currently neither APSIM-Oryza (nor its parent model ORYZA2000) simulates this phenomenon – crops can become submerged, yet happily continue to grow. Only one dataset exhibited submergence (#8), and we were able to adequately simulate the crop behaviour by creating a bootleg version of APSIM-Oryza in which the variables for crop_height and pond_depth were compared daily. When pond_depth > (crop_height*0.9), all simulated phenological development and biomass accumulation was put on hold. The applied empirical factor (0.9) was determined iteratively using comparisons between simulations and the observed data for both crop duration and biomass production. When the condition ceased to apply, crop growth was again continued with no damage component simulated.

3.7.8. AWD soil cracking

Soil cracks can extend below the plow-pan during drying phases in AWD rice irrigation or rainfed lowland rice, particularly when puddling has been performed and the drying period between flooding is long (ie the more extreme AWD applications, or long dry spells in rainfed cropping). When this occurs it may result in temporarily amplified saturated percolation rates on re-application of water, compared to continuously flooded treatments, before such time that the cracks re-seal (Bouman and Tuong, 2001; Tan et al., 2013). The result is temporarly increased water usage. Differences occur between soils of different clay content. APSIM currently does not simulate these dynamics in the saturated percolation rate ($K_s$).

3.7.9. Tillage impacts and soil property changes

When soils are puddled prior to the rice phase, then later tilled prior to the wheat phase in conventional rice-wheat, rice-maize and similar systems, significant inter-seasonal changes occur in soil properties (Gathala et al., 2011). Puddling reduces the effective $K_s$ and bulk density (BD) of affected layers, whereas tillage after the rice phase increases effective $K_s$ by breaking the plow-pan. BD of surface soil layers is decreased. These are significant input parameters in sensibly simulating rice-wheat system performance (Gaydon et al., 2012a), however APSIM does not currently simulate these parameter changes in response to specified tillage events. For the relevant simulated datasets in this study, we found it was necessary to employ APSIM-Manager to specify an increase in $K_s$ of the plow-pan layer by 100% (a doubling), and a decrease in BD by 5% following post-rice tillage, and apply the reverse change to these parameters upon puddling at the start of the subsequent rice phase (following findings of Gathala et al., 2011). Bund height (APSIM resetable parameter max_pond) was also reset to zero via APSIM-Manager on the date of rice field drainage (representing opening of the bunds), and then again set to the actual experimental bund-height on the occasion of bund establishment pre-crop. Once again, these are not simulated in APSIM, they are specified by the user. APSIM does however simulate the effect of tillage on soil roughness and therefore runoff and water retained on the soil surface and available for infiltration – by automatically reducing the USDA curve number, as per specified parameters. As an example in our Punjab rice-wheat simulations, we found it necessary to reduce curve number by 10 in the case of each discing, and by 5 for harrowing. The curve number was reset to the default curve number when at least 40 mm of water was added to soil (by rain or irrigation) to simulate the collapse/smoothing of a freshly cultivated soil surface as a result of saturation and rainfall impact (Balwinder-Singh et al., 2015a).

3.7.10. Direct-seeded rice germination and emergence

The APSIM-Oryza model currently simulates emergence of direct-seeded rice crops on the same day they are sown. There is no attempt to simulate the processes of germination/emergence
3.7.11. Radiation-use efficiency in diffuse light conditions

In APSIM, radiation use efficiency (RUE) is defined as a species-specific parameter, and by default does not change with radiation environments. In reality, RUE increases with the fraction of diffuse radiation (FDR). The decline in total radiation in most part of China, particularly in the North China Plain (NCP), may have led to reduction in crop potential yield, the increasing trend in FDR could have compensated the reduction to some extent. To account for the RUE change affected by the changing FDR, we developed a simple approach to modify RUE in APSIM-Wheat and APSIM-Maize modules for simulation of the wheat-maize double cropping rotation in the NCP. RUE was linearly linked to average seasonal FDR (Chen et al., 2010c; \( \text{RUE}_{\text{wheat}} = 0.78 \times \text{FDR} + 0.78; \text{RUE}_{\text{maize}} = 0.93 \times \text{FDR} + 1.03 \)), which doubled the RUE for wheat and increased RUE for maize by 90% when FDR increased from 0 to 1.0 based on Rodriguez and Sadras (2007) and Roechette et al. (1996).

3.7.12. Leaf growth

Initial simulations of winter wheat growth in North China Plain revealed that APSIM severely under-predicted LAI, biomass, and yield of wheat. Detailed process-level analysis indicated that the underestimation was mainly due to the incorrect temperature response of physiological processes implemented in the model for wheat. In the original model, it assumed that all the green leaves and the wheat plants are killed when daily minimum temperature drops to below \(-15\) °C. In the NCP, the winter wheat can survive at
a temperature as low as $-20\degree C$ (Jin et al., 1994) therefore as $-20\degree C$ was used instead of $-15\degree C$ as the threshold value in the model ($x_{\text{temp.senescence}}$). This change avoided the total winter killing of the wheat plant, but overestimation of early LAI and growth and underestimation of later LAI and biomass became apparent. Additional modifications were made to the temperature response of thermal time calculation and the temperature response of RUE for wheat based on Wang and Engel (1998) and Porter and Gawith (1999), which led to further improvement of LAI, biomass, and grain yield simulations for wheat (Chen et al., 2010b).

3.8. Datasets which we couldn’t simulate

Several categories of experiments were excluded from our assembled datasets on the basis of known APSIM model limitations.

3.8.1. Micronutrient response experiments

APSIM assumes that all crop nutrients apart from N and P are non-limiting. This assumption stems from its Australian heritage, where farmers routinely fertilize adequately to avoid micronutrient constraints. However in low-input cropping systems of Asia the limitations placed by shortages (or excesses) in micro-nutrients (and consequently how to manage them within local resource constraints) is sometimes a relevant research issue. The current inability of APSIM to quantify crop responses to micronutrient stress, precluded the use of experimental datasets exploring these issues.

3.8.2. Greenhouse gas (GHG) emissions in rice cropping

APSIM has been successfully used in studies of GHG emissions in non-flooded cropping systems (wheat-sorghum, Huth et al., 2010; sugarcane, Thorburn et al., 2010), however the model currently cannot simulate GHG emissions in flooded systems under anaerobic conditions. Such datasets were excluded.

3.8.3. Salinity response in non-rice crops

Datasets involving non-rice crops and salinity response were excluded from consideration as the majority of APSIM crops do not yet have the capacity to simulate crop salinity stress.

Fig. 16. Comparison between observed and simulated soil organic carbon (t ha$^{-1}$) for the experiments of Wang et al., 2014. (dataset # 27). A mixture of wheat and wheat-maize cropping systems were examined over four (4) Chinese locations, over a period of 20+ years with a range of fertiliser and stubble treatments. NPK = application of compound inorganic fertilizers; NPKSt = application of inorganic fertilizers and stubble retention; CK = control. Symbols show the observed values and lines show the simulated values.

Fig. 17. APSIM-Oryza simulated versus observed rice production (Mg ha$^{-1}$) for both FACE and ambient CO$_2$ treatments: (a.) Yang et al., 2006, 2008 (dataset # 25); (b.) Kim et al., 2003 (dataset # 24). Comparison for medium N treatments shown. No replicate variability information (error bars) available.
4. Discussion

4.1. Model performance

When evaluating model performance and determining whether ‘acceptable’ performance has been achieved, a powerful technique is to demonstrate that the RMSE between simulated and observed values is around the same quantum (and ideally smaller) than the standard deviation within the observed values (for example, across experimental replicates). When this is true, it essentially demonstrates that the model can simulate the observed behaviour within the bounds of experimental uncertainty. In reality, this is all you can ever expect a model to do. It is unrealistic to expect modelled results to be perfectly the same as the average of the observed values, because as we know (via experimental replicates) there is uncertainty in the observed data. Being exactly the same as the average provides no additional meaning above ‘being within the range of experimental uncertainty’, in this context. It is also arguable whether it is valid to state that one model is performing better than another model on the basis of a lower RMSE, when both model RMSE’s are within the range of the experimental uncertainty.

Model evaluation has taken place over a broad spectrum of Asian locations, environments, and crop management practices, with a focus on APSIM’s ability to simulate individual crops as well as wider cropping system performance. The model performed well in simulating above-ground biomass and grain yield across the diversity of our assembled datasets.

Good quality datasets for soil water dynamics and water-balance components are relatively rare in Asia. For potted rice (dataset #5, CS and AWD, Sudhir-Yadav et al., 2011a,b), non-flooded maize and wheat crops (dataset #38, rainfed and irrigated wheat, Balwinder-Singh et al., 2011; dataset #28, rainfed and irrigated wheat and maize, Chen et al., 2010a) acceptable performance was demonstrated by APSIM in simulating soil water balance in individual soil layers and system water-balance components (input requirement (irrigation), transpiration, evaporation, runoff and drainage). These were achieved in conjunction with sensible simulation of crop production (Figs. 12–16), often considered to be a key criterion of “getting the right results for the right reasons” (Gaydon et al., 2012a; both above and below-ground processes correct). Simulation of soil moisture and system water balance using APSIM has been more thoroughly evaluated outside of Southern Asia (for example, in Australia; Verburg and Bond, 2003), where greater rigour is possible due to the availability of quality datasets. Given the added confidence from studies such as these, and the lack of any major differences between soils of Asia and the diversity of soils in Australia, we propose confidence in APSIM’s capacity to simulate the system water balance components and soil moisture dynamics.

Confidence in a model’s capacity to respond sensibly in simulating crop response to changes in CO2 is a pre-requisite to its successful employment in climate change studies, either on impacts or adaptations. We used FACE datasets for rice from Japan and China for this evaluation. On the basis of this limited available data, we conclude that APSIM is capable of sensibly simulating CO2 response of rice crops in the region under non-stressed water conditions (Fig. 17) and across a range of N-stress conditions imposed in these FACE experiments. Further data evaluating potential interaction between water stress, nitrogen, and CO2 would provide enhanced confidence in model performance, particularly for simulation of rainfed rice systems in future climates. We have reflected on the use of APSIM with FACE experiments in other parts of the world to give confidence in CO2 response of wheat, maize and other crops (Asseng et al., 2004; O’Leary et al., 2015). APSIM has been successfully used in a number of published climate change studies in Asia (Bhopal, India (Mohanty et al., 2015); Upper Gangetic Plains, India (Subash et al., 2014a); and the North China Plain (Zhang et al., 2013)).

Although this evaluation illustrated good model performance, in some situations that was achieved by overcoming significant input parameter estimation challenges. Even though these were surmounted via ‘work-arounds’ described earlier, they may in reality indicate model deficiencies and the need for future model improvement to make this process less challenging to inexperienced model users. These are discussed below. Some relate to improvements required in external models which plug-in to the framework (rice, for example; APSIM-Oryza/ORYZA2000) rather than deficiencies within APSIM itself.

4.2. Areas indicated for model improvement

When a model gives an erroneous estimation for crop yield (for example), it is naive to immediately assume that the problem lies with the crop component of the model. Model performance is equally dependant on accurately representing the resource supply terms, such as soil nutrients, water and climate (Carberry et al., 2009). In presenting the following suggestions that relate to crop model improvement, we have implicitly assumed accurate specification and simulation of supply terms (soil, climate) in the relevant simulations.

4.2.1. Crop phenology

Simulation of crop phenological development is closely tied to success in simulating grain yields. The work-arounds detailed in the previous section facilitated this performance in certain situations, and we refer to these in recommending the following areas for model improvement.

- **Rice N-stress response:** In the datasets (n=5) in which crop N-stress was extreme there was clear indication that further improvement of the APSIM-Oryza phenological model is required. This situation arises from the original phenological model within APSIM-Oryza and ORYZA2000 (Bouman and van Laar, 2006) developed using data from non-N-stressed crops, not capturing the effects of N-stress on phenology. Improvement in simulation of this aspect is indicated, however relevantly for this analysis the subsequent simulation errors for both rice grain yield and biomass were within the bounds of experimental uncertainty (standard deviation of observed data) indicating model performance is still acceptable. Incorporation of an N-stress factor (0–1) applied to phenological development could represent one potential avenue (as used in other APSIM crops), and we suspect this factor may have different effects at different crop stages (for example, abiotic stresses may speed up crop development between certain stages while slowing it down between others (Angus and Moncur, 1977)). There may also be interactions with other crop stresses (temperature, water etc.)

- **Response to submergence:** The ability simulate rice crop responses to occasional submergence is of growing relevance to modellers of rice-cropping in the major rice-growing deltas of South and Southeast Asia (Ganges-Brahmaputra, Irrrawaddy, Chao Prahya, Mekong) threatened by rising seas levels and a changing climate. We realise the work-around described in Section 3.7.5, whilst conceptually meaningful and adequate for the single submergence dataset we simulated, may be too simplistic to apply to wider situations. Deeper investigation of this issue is warranted using a broader range of experimental submergence datasets. We recommend the development of generic algorithms describing submergence damage to rice crops as a function of time, crop stage, submergence depth and possibly other variables, and subsequent incorporation into APSIM-Oryza and similar rice models.
• **Response to mulching:** Mulched wheat crops in north-west India exhibited delayed phenology in comparison with non-mulched crops (Balwinder-Singh et al., 2011; Sidhu et al., 2007). Currently there is little published research to indicate whether the primary driver for this phenomenon is canopy or soil temperature differences resulting from the presence of the mulch. Both canopy temperature and soil temperature may require simulation – at early crop stages, soil temperature may be more appropriate because the growing point is close to the soil, while at later stages canopy temperature will dominate, particularly under dry and irrigated conditions. While soil temperature is currently simulated in APSIM, there is no canopy temperature module at the moment. We suggest that the current APSIM pheno-logy routines for all crops may require modification to employ some combination of canopy and soil temperature, rather than ambient temperature (as per current approach). Of course, this implicitly also requires sensible simulation of the mulch effects on canopy temperature and soil temperature.

• **Response to high temperatures:** Our simulation exercise demonstrated that high temperature crop development responses can effectively be simulated in APSIM using a modified thermal time accumulation function (Section 3.7.2). The veracity of this approach has been alluded to by other authors for high temperature cropping environments (Lobell et al., 2012). We suggest that this modified relationship should be considered for incorporation into the APSIM standard release, with all crops being addressed. It should also be a varietal input characteristic, as different varieties exhibit different heat tolerances and development responses to both high and low temperatures. In general, the temperature response simulations of key physiological processes need to be revisited. This includes temperature response of not only phenology, but leaf growth, RUE and other processes for different crops, and whether soil/canopy temperatures should be used. In addition, the impact of extreme temperatures has been poorly defined and rarely tested, including both low and high temperature impacts. For example, high temperature can impact sterility and grain abortion, in addition to adversely impacting other processes. The conceptual model of Barlow et al. (2015) may be helpful.

• **Simulation of emergence in direct-seeded rice:** The APSIM-Oryza model currently simulates emergence of the rice crop on the same day it is sown. We suggest incorporation of algorithms to simulate the process of germination and emergence as a function of a thermal time target, depth of sowing, and soil moisture content (like already used in the other APSIM crop modules – Wang et al., 2002). This aspect is particularly relevant in Asia for simulation of tight cropping sequences which include direct-seeded rice in intensive cropping patterns with other crops, and timeliness is important.

• **Leaf Growth:** Leaf growth and canopy development are simulated differently in various APSIM crop models. Model evaluation using our Chinese datasets revealed that APSIM underestimates the impact of crop density on crop growth for both wheat and canola (Zhang et al., 2012). A leaf cohort approach has been developed for APSIM-Sorghum (Van Oosterom et al., 2001) and is currently being developed for APSIM-wheat and other crops. This approach aims to capture the impact of plant density and environmental conditions on tiller and leaf dynamics, thus to enable better simulation of leaf area growth and the impact of plant density.

• **Salinity response in non-rice crops:** Radanielson et al. (2016) report successful modifications to the APSIM-Oryza model, allowing simulation of rice crop response to changing levels of soil salinity. The model now characterises rice response in terms of varietal tolerance and resilience, with osmotic stress and sodium ion-toxicity stress both affecting simulated processes including photosynthesis, stomatal conductance, and transpiration. APSIM-Wheat has the capacity to respond to saline conditions via simulated reduction in soil PAWC (plant available water content) (Hochman et al., 2007), however this simplified approach is untested in Asian conditions to the best of our knowledge. Managing salinity in Asian cropping systems is likely to be of growing relevance to the major rice-growing deltas of South and Southeast Asia (Ganges–Brahmaputra, Irrawaddy, Chao Phraya, Mekong) threatened by rising sea levels and a changing climate. We suggest evaluation of the APSIM-Wheat approach using available salinity datasets in these regions, and a subsequent enhancement of all relevant APSIM crops – either using this approach or the more physiologically-detailed approach of Radanielson et al. (2016).

• **Floret sterility in rice due to high temperatures:** APSIM-Oryza currently calculates rice floret sterility based on average ambient maximum temperatures during the floral development and flowering period. We suggest modifying the model to simulate crop canopy temperature directly, and subsequently calculate sterility on that basis (to automatically simulate the process which we performed manually, as described in 3.7.3). This same requirement applies to all APSIM crops.

**4.2.3. Soil processes**

• **Tillage impacts:** APSIM currently simulates tillage effect on USDA curve number (runoff-infiltration ratio) and soil evaporation (as well as other parameters; Connolly et al., 2002), but requires manual resetting of some parameters describing other affected soil properties in rice-wheat and similar cropping systems with transitional puddled and tilled phases (soil bulk density and saturated percolation rate; Section 3.7.7). We judge there are now sufficient data available from published experiments investigating CA effects on soil properties to improve the APSIM-Soilwat model in automatically modifying such parameters when tillages are specified.

• **GHG emissions:** There is a clear development need for sensible APSIM GHG simulation capacity in rice-based systems with seasonal flooding, building on the successful capacity already added for non-flooded cropping phases (CO₂ and N₂O; Thorburn et al., 2010; Huth et al., 2010). Further development work to segregate C and N emissions into specific pools (N₂, NO₃, N₂O, CO₂, CH₄) under anaerobic soil conditions is suggested.

• **Micronutrient dynamics:** APSIM’s primary development heritage is in dryland cropping systems of Australia, where growth-limiting deficiencies in soil micronutrients (S, Zn, Ca, Mg etc.) are routinely removed by farmer fertilizer applications. Consequently APSIM crops currently assume micronutrients are non-limiting in simulations. In low-input cropping systems, this
is often not the case and research questions exist around how to best manage deficiencies in micronutrients in different soils, climates, and cropping systems (for example, to encourage zinc fortification in rice – Phattarakul et al., 2012; Rose et al., 2013). Toxicity of some micronutrients also represent researchable management issues (Suriyagoda et al., 2016) which simulation models currently cannot serve. We suggest that consideration is given to establishing micronutrient stress factors in APSIM crop modules, similar to the way N and P-stresses are simulated. This would require simultaneous development of APSIM soil modules to simulate the dynamics of micronutrients over depth and time, as a function of imposed farmer management.

4.3. General comments on model applicability and limitations

Many of the looming research questions for the south Asian region relate to development of best management practices for cropping systems within their local environmental and socio-economic constraints – aiming to maximise land and/or water productivity whilst minimising negative environmental outcomes (Godfray et al., 2010; Gathala et al., 2014). In this regard we have demonstrated APSIM’s capacity to reliably simulate sequences of crops and system dynamics over a broad expanse of Asian agriculture, in addition to gaining some confidence in the model’s response to increasing CO₂, the effect of stubble, deficit irrigation, increased cropping intensity, and fertiliser practices. Elsewhere, the APSIM model has been demonstrated as particularly strong in its capacity to simulate detailed farmer management practices and decision-trees (Holzworth et al., 2014; Amarasingha et al., 2015; Balwinder-Singh et al., 2015a,b; Hochman et al., 2017a,b), allowing simulations to reflect how farmers will actually respond in different seasons and conditions. This makes APSIM a robust choice for a cropping systems model when the research questions centre on development of detailed farmer adaptation strategies to external forces of change. To the best of our knowledge, no model apart from APSIM currently has the ability to simulate detailed farmer management strategies in cropping systems, which reflect year-to-year changes in actual on-ground farmer actions in response to prevailing environmental conditions. Our analysis which encompassed 12 countries, 43 datasets, 7 crop types and numerous crop rotations, adds to previously more disparate crop model evaluation exercises performed for other models in the region (for example, DSSAT – Timsina and Humphreys 2006a,b: 11 datasets, primarily rice and wheat only as individual crops rather than systems (only 1 sequence dataset which performed poorly), crop components only, 7 countries; INFOCROP – Aggarwal et al., 2006b: 11 datasets, rice-wheat systems only, all experiments at New Delhi, India). The demonstrated performance of APSIM in simulating cropping performance (crop and soil dynamics) with and without mulch position it strongly to contribute to future research in Conservation Agriculture in Asia, as does its demonstrated ability to robustly simulate cropping sequences and the impacts of associated cropping systems intensification. Simulations of diverse cropping sequences up to seven consecutive crops showed no trend of increasing residual error between simulated and observed yields, despite no re-setting of soil parameters (water, nutrients) between crops, suggesting robust simulation of long-term system processes. As far as we are aware, our method of analysing residual error with advancing crop number to evaluate the performance of a cropping sequence model, is novel.

Having a model which is ‘capable’ is clearly essential, however correct parameterisation, calibration and validation procedures will always be important for APSIM users in obtaining robust model performance at a local scale. This analysis has highlighted numerous strengths, however has also identified some aspects which make APSIM parameter estimation challenging or else limit APSIM’s applicability. Relevant model improvements have been suggested.

5. Conclusions

This research shows that APSIM performed in a statistically robust manner in simulating cropping system performance over a wide range of crops, varieties, environments and management practices in Asia. Aspects evaluated include production and phenology of individual crops (rice, wheat, maize, others); the ability to robustly simulate soil water and nutrient dynamics under cropping sequences without soil variable resets, and crop response to CO₂. For rice, the major crop of the study region, our analysis indicates that APSIM performs better in simulating PRT than DSR, and better in simulating rice production under continuously-flooded water management than either AWD irrigation or rainfed (no irrigation) systems. However simulation of each of these important systems indicated no significant difference between simulated and observed rice grain yields at 95% confidence level, hence APSIM’s performance across the systems can still be categorized as acceptable, and within the range of experimental data uncertainty. Similarly, performance in simulating the other major crops of the region (notably maize, wheat and cotton) were shown to be within the bounds of experimental error. Input parameter estimation challenges were encountered, and although ‘work-arounds’ were developed and described, in some cases these actually illustrate model deficiencies which need to be addressed. Desirable future improvements have been identified to better position APSIM as a useful tool for Asian cropping systems research into the future. These include aspects related to harsh environments (high temperatures, salinity, rice crop submergence), conservation agriculture, greenhouse gas emissions in rice systems, as well as aspects more specific to Southern Asia and low input systems (such as deficiencies in soil micro-nutrients).

Acknowledgements

The authors would like to thank the Australian Department of Foreign Affairs and Trade (DFAT, formerly AusAID), and the Australian Centre for International Agricultural Research (ACIAR) for jointly funding the underpinning research projects supporting this work. We gratefully acknowledge the China Scholarship Council for providing PhD and visiting scholarships under the CSIRO—Chinese Ministry of Education (MoE) PhD Fellowship Research Program to support the modelling work in China. We also express our appreciation to Susanna Bucher (via Achim Dobermann), Annie Boling (via Bas Bouman), and Sudhir Yadav for provision of unpublished datasets for model testing.

References


Hochman, Z., Prestwidge, D., Carberry, P.S., 2014. Crop sequences in Australia’s northern grain belt have historically yielded more grain and were more efficiently managed than implied by the sum of their parts. Agric. Syst. 129, 124–132.


Rai, G.S., 2014 Chapter 4- Application of APSIM in evaluation of drought tolerance variability of rice for the low rainfall conditions of the semi-arid SAARC–Australia Project – developing capacity in cropping systems modelling for South Asia (Gaydon, Sanyed, Roth, editors), SAARC Agriculture Centre Monograph, SAARC Agricultural Centre, Karachi, Pakistan, Farm Gate, Dhaka-1215, Bangladesh, 259 pages ISBN: 978-984-33-7469-1.


