Evaluation of the APSIM common bean model using different cultivars and water-management scenarios

Seyedreza Amiri[1*], Nader Zakeri[2] and Tayebeh Yousefi[3]

[1] *Department of Production Engineering and Plant Genetics, Faculty of Agriculture, Higher Education Complex of Saravan, PO Box 9951634145, Saravan, Iran*
[2] *Faculty of Agriculture, Shahid Bahonar University, Kerman, Iran*
[3] *Higher Education Complex of Saravan, Saravan, Iran*
* Corresponding author, E-mail: amiriseyedreza86@gmail.com or r.amiri@saravan.ac.ir

Abstract

The common bean (*Phaseolus vulgaris*) is a widely consumed legume worldwide and holds significant value in terms of direct human consumption, surpassing all other legume crops combined. This study aimed to assess the applicability of the APSIM common bean model under different water-management conditions and cultivars in southeastern Iran. A two-year field experiment was conducted at the Agricultural Research Station of Shahid Bahonar University of Kerman, Iran, spanning from 2020 to 2021. The experiment followed a split-plot design with a randomized complete block structure and included four replications. The primary irrigation factor consisted of three levels, representing 80%, 60%, and 40% of crop capacity, while the secondary factor encompassed early and late-maturity cultivars. To evaluate the model's performance, simulated and measured grain yield, total dry matter, leaf area index (LAI), and soil water content were compared using the adjusted coefficient of correlation, normalized root mean square errors (nRMSE), and model efficiency (EF). The results indicated a satisfactory agreement between predicted and observed grain yield (nRMSE = 12% and EF = 0.92). Similarly, the agreement between simulated and observed total dry matter was reasonable (nRMSE = 13% and EF = 0.83). The observed and predicted soil water content also exhibited good agreement. Consequently, the APSIM common bean model proves suitable for research purposes, particularly in the areas of irrigation and cultivar selection.

Introduction

P. vulgaris, a plant species extensively grown and cultivated, is a member of the Fabaceae family, commonly known as the legume family. This particular species holds great importance as a staple food crop worldwide, contributing significantly to the provision of protein, fiber, vitamins, and minerals. Furthermore, the common bean serves as a vital source of dietary protein for numerous individuals residing in South and Southeast Asia, a considerable portion of which adhere to vegetarian diets. With approximately 24% of easily digestible protein, and an abundance of fiber, antioxidants, and phytonutrients, common bean seeds are consumed in various forms, such as whole or split, ground into flour, or utilized as sprouts[1]. Beyond their nutritional value, common beans play a pivotal role in the realm of sustainable agriculture. Their unique ability to fix nitrogen in the soil not only enriches its fertility but also lessens reliance on artificial fertilizers. Crop simulation models have been utilized to examine the impacts of the most effective management practices on crop productivity over the previous two decades. To illustrate, the CROPGRO model[2] can simulate the growth and development of three legume crops, namely common bean (*Phaseolus vulgaris*), peanut (*Arachis hypogaea*), and soybean (*Glycine max*). Additional generic models encompass STICS[3] and EPIC[4]. These models are comprised of a singular collection of subroutines and are parameterized with coefficients that are externally sourced to the code. Nevertheless, it has been recognized that this particular approach, while enhancing the functionality of the model for a variety of species, may result in a trade-off in terms of physiological rigor and predictive capability[5] .

Different approaches are used in current models for legume production to simulate growth and development. One simpler approach, as exemplified by Sinclair's soybean model[6] , has been expanded to include cowpea (*Vigna unguiculata*), black gram (*Vigna mungo*)[7] and chickpea (*Cicer arietinum* L.)[8]. In contrast, the CROPGRO model provides a more detailed coverage of various physiological processes. Both models incorporate the dynamics of biomass and nitrogen in response to climate, nitrogen, and water supply. However, CROPGRO necessitates the specification of more than 40 crop-specific and cultivar-specific parameters, which can complicate the interpretation and limit its adaptability to different environments and cultivars. The APSIM-legume model seeks to strike a balance between the two by offering a comprehensive simulation of crop growth and development without the need for a large number of parameters that are difficult to quantify[9].

The APSIM cropping systems framework, as demonstrated in the works of Keating et al.[10] and Holzworth et al.[9], is a model that has a proven record in accurately modeling the performance of various cropping systems, including rotations, fallowing, and the dynamics of crops and the environment[11−13]. What sets APSIM apart from other crop models is its unique approach of primarily focusing on simulating the supply of resources for crops, rather than solely focusing on the demand for resources. In this framework, the soil plays a central role in

the simulation, as crops, influenced by weather and management practices, interact with the soil in one state and leave it in a different state for the subsequent crop[14]. The aim of this particular investigation, however, was to assess the effectiveness of APSIM-legume in simulating the growth and development of two different cultivars of common bean, in relation to various water management scenarios.

Materials and methods

Model description

The legume module, much like other crop modules within APSIM emulates the growth, development, yield, and nitrogen accumulation of legumes in response to various environmental factors such as temperature, radiation, photoperiod, soil water, and nitrogen availability. The principles employed in APSIM-legume have been expanded upon from the original modeling approaches introduced by Ritchie[15] and subsequently summarized by Ritchie[16]. The model operates on a daily time-step and has been specifically designed to simulate a homogeneous field, predicting grain yield, crop biomass, crop nitrogen absorption and fixation, as well as the distribution of resources within the plant on a spatial basis. The methodologies employed in modeling crop processes aim to strike a balance between providing a comprehensive depiction of the observed variations in crop performance across diverse production environments and minimizing the number of parameters that are difficult to quantify[17].

When parameterizing a generic model for legume species, particularly the common bean, it proves beneficial to introduce the concept of essential, desirable, and optional parameters. Essential parameters are those that exhibit the greatest sensitivity within the model and often differ significantly across various species. Consequently, these parameters must be determined through experimentation for each species[18]. Examples of essential parameters include the phyllochron, radiation extinction coefficient, and radiation-use efficiency. While it is sometimes possible to define essential parameters based on published sources, due to their critical nature, it is often essential to verify the value of such parameters through local experimentation using commonly cultivated varieties. The

development of the common bean module can be attributed to Peter Carberry and Michael Robertson, and their work is elaborated upon in the paper authored by Robertson et al.[19].

Calibration and validation of the common bean-legume model

The calibration of the model was performed by utilizing measured data obtained from a two-year field experiment that was carried out during the growing seasons of 2020 and 2021. The experimental design employed was a split plot, based on a randomized complete block design with four replications, and it was conducted at the Agricultural Research Station of Shahid Bahonar University of Kerman in Iran, located at the coordinates 32.38° N and 51.40° E. The primary factor considered in the experiment was irrigation, which consisted of three levels: 80%, 60%, and 40% of the field capacity. Additionally, the sub-factor taken into account was the maturity stage, with early and late-maturity being the respective categories.

The date of planting was the 10th of May for both years. Each experimental plot consisted of a length of 6 m and four rows with a spacing of 50 cm between rows, and a plant density of 30 plants per square meter. The soil texture was classified as silty loam. Irrigation was carried out using a drip system. The amount of water applied to each plot was measured using a counter. All plots were irrigated uniformly every four days until the four-leaf stage. After this stage, irrigation was adjusted based on the field capacity using the weight method. The common bean crop was harvested on the 10th of July and the 1st of July in the years 2020 and 2021, respectively. To determine the grain yield, the final harvest involved the two middle rows (10 m^2) at physiological maturity (with a humidity level of 12%). The field data were used to estimate genotype-specific parameters for the model (as shown in Table 1). The model was calibrated using a trial and error approach, aiming to minimize the difference between observed and simulated values. Parameters that had a significant impact on dry matter and leaf area index (LAI) were adjusted accordingly. This process was done until the model's simulated values closely matched the observed values for all treatments[20].

To calibrate and validate the water content, soil moisture content data collected from two field experiments conducted in 2020 and 2021 were utilized. To validate the crop model, the

Table 1. Cultivar-specific parameters for a early- and late-maturity common bean cultivars.

Parameter	Units	Parameter description	Late-maturity cultivar	Early-maturity cultivar
Cardinal temperatures for thermal time calculation	°C	Base temperature	7.5	7.5
		Optimum temperature	30	30
		Maximum temperature	40	40
Shoot_rate	Degree-days/mm	Thermal time required per mm elongation by young shoot before emergence	0.6	0.56
node_app_rate	Degree-days	Thermal time required for node appearance on main stem	100	92
Leaves_per_node	lf/node	No. of leaves per plant per main stem node	2	2
Extinction_coef		Extinction coefficient (at default row spacing)	0.40	0.40
Rue	g/MJ	Radiation-use efficiency	0.94	0.94
frac_leaf_pre_flower		Fraction allocated to leaves pre-flowering	0.55	0.60
frac_leaf_grain_fill		Fraction allocated to leaves in grain fill	0.30	0.35
frac_stem2 pod		Fraction allocated to pod before grain fill	0.46	0.49
frac_pod2 grain		Fraction allocated to pod relative to grain during grain fill	0.28	0.28
n_conc_crit	g/g	Critical nitrogen concentration of grain	0.045	0.045
Specific_root_length	mm/g	Specific root length	65000	65000
Trans_eff_coef	Pa	Transpiration efficiency coefficient	0.0055	0.0055

dry matter, LAI, and grain yield traits acquired from these field experiments were taken into account. To assess the variances between the observed and simulated data, various metrics were employed[21].

These metrics included the coefficient of determination (R^2), the 1:1 line, the normalized root mean square error (nRMSE) as proposed by Wallach & Goffinet[22], and the model efficiency (EF) as defined by Willmott[23]. These metrics were utilized to compare the observed and simulated values of key parameters such as grain yield, dry matter, leaf area index (LAI), and soil moisture content.

$$nRMSE \ (\%) = \sqrt{\frac{\sum_{i=1}^{n}(S_i - O_i)^2}{n}} \times \frac{100}{\overline{O}} \qquad (1)$$

$$EF = 1 - \frac{\sum_{i=1}^{n}(S_i - O_i)^2}{\sum_{i=1}^{n}(\overline{o} - O_i)^2} \qquad (2)$$

Where, $\overline{O}$, O, S, and n represent the average observed data, observed data, simulated data, and number of observations, respectively. The accuracy of the model improves as the nRMSE value approaches zero. Additionally, a value of one for EF (model efficiency) indicates that the simulated values better depict the trend in the measured data compared to the average of the observations[24]. Furthermore, a higher R^2 value nearing unity demonstrates that the model accurately replicates reality.

Results and discussion

Model parameterization

For the majority of the plant's growth period, there was a notable correspondence between the leaf area index and the distribution of dry matter among different plant tissues, in accordance with the observations made in the experimental study of two distinct cultivars. The calibration of the model employed in this study, as demonstrated in Fig. 1, yielded results that reasonably approximated the predicted values of dry matter for both the early-maturity and late-maturity cultivars, based on the 2020 experimental data. The calibration analysis revealed that the nRMSE and EF values for dry matter were 7% and 0.86, respectively, for the early-maturity cultivar, and 9% and 0.97, respectively, for the late-maturity cultivar. Moreover, the coefficient of determination (R^2) values obtained from the regression analysis, comparing the observed and simulated values of dry matter ranged from 0.92 to 0.96 for the two cultivars. Furthermore, when it came to the leaf area index, the RMSE and EF values were 12% and 0.96, respectively, for the early-maturity cultivar, and 10% and 0.94, respectively, for the late-maturity cultivar, as depicted in Fig. 2.

Model evaluation

Figure 3 depicts the outcomes of model validation for two varieties of early- and late-maturity using experimental data collected in 2021. The validation results demonstrated that the model accurately predicted the leaf area index (LAI) (Fig. 3). For example, the normalized root mean square error (nRMSE) and model efficiency (EF) for LAI were found to be 20% and 0.62, respectively. Furthermore, the coefficient of determination (R^2) value for the regression analysis between the observed and simulated LAI was determined to be 0.83 for both cultivars.

Moreover, the model validation results indicated that the model also effectively predicted the dry matter and grain yield of common beans in 2020 and 2021. According to Fig. 3, the values of nRMSE, EF, and R^2 for grain yield were 12%, 0.92, and 0.97, respectively. Similarly, for dry matter, these values were 13%, 0.83, and 0.91, respectively. These findings suggest that the APSIM-legume model was capable of accurately simulating
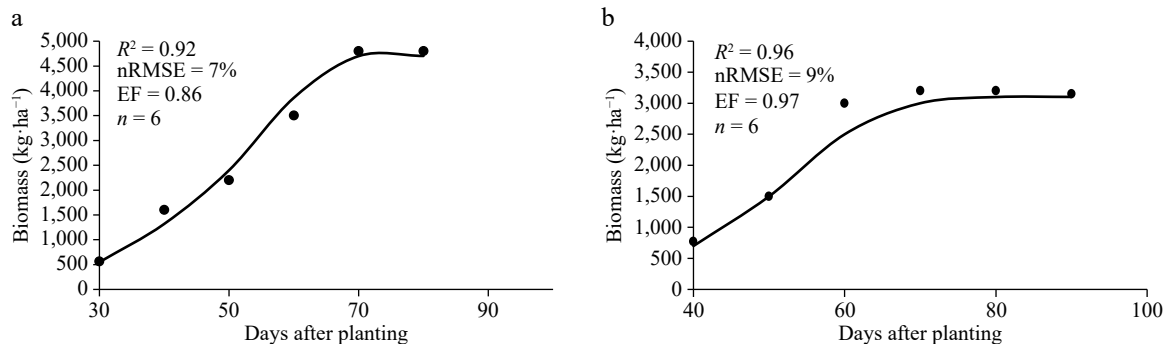
Fig. 1 Simulated (lines) and measured (data points) values of dry matter of (a) early- and (b) late-maturity cultivars.
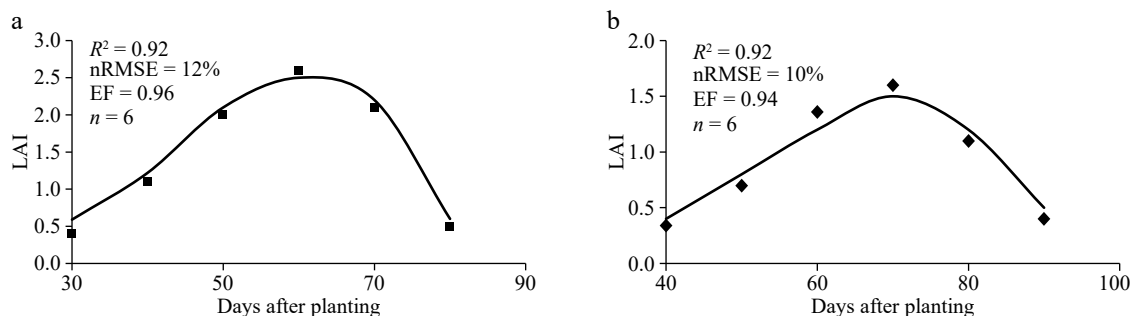
Fig. 2 Simulated (lines) and measured values (data points) of leaf area index (LAI) of (a) early- and (b) late-maturity cultivars.
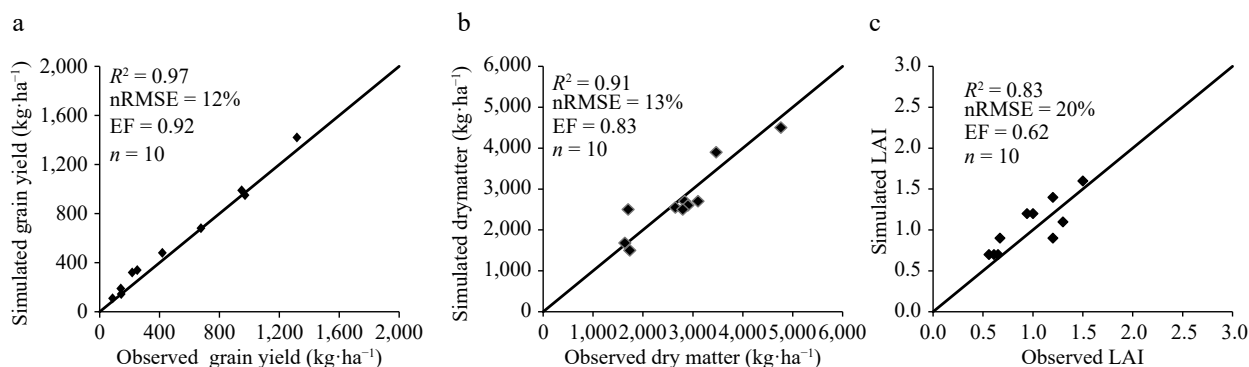
Fig. 3 (a) Simulated vs measured grain yield, (b) dry matter and (c) maximum LAI of common bean in different treatments (2020 and 2021 field experiments). Continuous line : 1 to 1 line.

the growth and grain yield of common beans for two consecutive years and across different cultivars.

The nRMSE values for soil water content exhibited a range of 14% in the year 2020 to 12% in the year 2021, as depicted in Fig. 4. Additionally, the EF values displayed variation, with a range of 0.94 in 2015 to 0.80 in 2016, also shown in Fig. 4. This provides evidence that the APSIM-legume model accurately captured the soil water balance during the growth period of the common bean across different years. The APSIM-legume model underwent testing using an independent set of experiments, primarily conducted in the tropics and subtropics of Australia. These experiments encompassed various factors such as cultivar, sowing date, water regime (irrigated or dryland), row spacing, and plant population density[19]. The model, introduced by Robertson et al. in 2002[19], aimed to simulate crop growth and development with comprehensive coverage, eliminating the need for an excessive number of parameters. The model adopted a generic approach, acknowledging the shared physiology and simulation methods across numerous legume species. When simulating grain yield, the model accounted for 77%, 81%, and 70% of the variance, with RMSE values of 31, 98, and 46 $g \cdot m^{-2}$ for common bean (n = 40, observed mean = 123 $g \cdot m^{-2}$)[25]. However, the simulation of biomass at maturity proved to be less accurate, explaining 64%, 76%, and 71% of the variance, with RMSE values of 134, 236, and 125 $g \cdot m^{-2}$ for common bean, peanut, and chickpea, respectively.

However, it has been asserted that the occurrence of a 10%–20% normalized root mean square error (nRMSE) threshold in the context of model inaccuracies is quite prevalent[19,26]. Indeed, attaining model errors below these thresholds is exceedingly challenging because models are benchmarked against experimental data, which inherently entails a degree of error. Furthermore, the incorporation of observation errors in the inputted parameters and variables further complicates the achievement of lower error levels.

Conclusions

The present experiments showed that the model's simulated values of LAI, dry matter, grain yield, and soil water content were in general agreement with the observed values across different treatments. This indicates that the model has a robust predictive capability. It can therefore be concluded that the model can be used to analyze crop yield and its limitations in response to environmental conditions and management input.
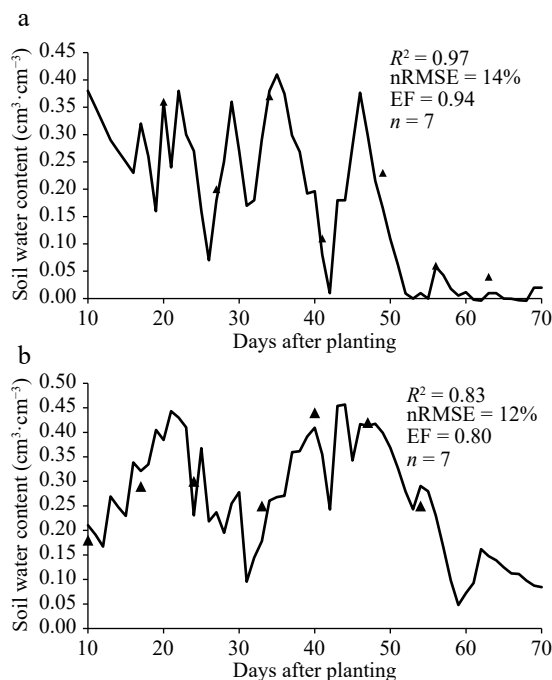
Author contributions

The authors confirm contribution to the paper as follows: methodology, evaluation, writing – original draft: Amiri S; writing – review & editing: Zakeri N, Yousefi T. All authors reviewed the results and approved the final version of the manuscript.

Data availability

Data will be made available upon reasonable request.

Acknowledgments

Fig. 4 Comparison of simulated (lines) and observed (data points) soil moisture content at the top soil layer (0–20 cm) in two growing seasons: (a) 2020 and (b) 2021.

Province for their assistance in conducting the present research.

References

1. Tang D, Dong Y, Ren H, Li L, He C. 2014. A review of phytochemistry, metabolite changes, and medicinal uses of the common food mung bean and its sprouts (*Vigna radiata*). *Chemistry Central Journal* 8:4
2. Boote K J, Jones JW, Hoogenboom G. 1998. Simulation of crop growth: CROPGRO model. In *Agricultural Systems Modeling and Simulation*, eds. Peart RM, Curr RB. Boca Raton: CRC Press. pp. 651–92. https://doi.org/10.1201/9781482269765
3. Brisson N, Mary B, Ripoche D, Jeuffroy MH, Ruget F, et al. 1998. STICS: a generic model for simulation of crops and their water and nitrogen balances. 1. Theory and parameterisation applied to wheat and corn. *Agronomie* 18:311–46
4. Williams JR, Jones CA, Kiniry JR, Spanel DA. 1989. The EPIC crop growth model. *Transactions of the American Society of Agricultural Engineers* 32:497–511
5. Telen D, Logist F, Derlinden EV, Impe JF. 2013. On the trade-off between experimental effort and information content in optimal experimental design for calibrating a predictive microbiology model. *Journal de la Société Franç aise de Statistique* 154(3):95–112
6. Sinclair T R. 1986. Water and nitrogen limitations in soybean grain production I. Model development. *Field Crops Research* 15:125–41
7. Sinclair TR, Muchow RC, Ludlow MM, Leach GJ, Lawn RJ, et al. 1987. Field and model analysis of the effect of water deficit on carbon and nitrogen accumulation by soybean, cowpea and black gram. *Field Crops Research* 17:121–40
8. Soltani A, Sinclair T R. 2011. A simple model for chickpea development, growth and yield. *Field Crops Research* 124:252–60
9. Holzworth DP, Huth NI, de Voil PG, Zurcher EJ, Herrmann NI, et al. 2014. APSIM - evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software* 62:327–50
10. Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, et al. 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* 18:267–88
11. Carberry PS, Probert ME, Dimes J P, Keating B A. McCown R L. , 2002. Role of modelling in improving nutrient efficiency in cropping systems. In *Food Security in Nutrient-Stressed Environments: Exploiting Plants' Genetic Capabilities. Developments in Plant and Soil Sciences*, ed. Adu-Gyamfi JJ. vol 95. Dordrecht: Springer. https://doi.org/10.1007/978-94-017-1570-6_34
12. Whitbread AM, Robertson MJ, Carberry PS, Dimes JP. 2010. How farming systems simulation can aid the development of more sustainable smallholder farming systems in southern Africa. *European Journal of Agronomy* 32(1):52–58
13. Hochman Z, Prestwidge D, Carberry PS. 2014. Crop sequences in Australia's northern grain zone are less agronomically efficient than implied by the sum of their parts. *Agricultural Systems* 129:124–32
14. McCown RL, Hammer GL, Hargreaves JNG, Holzworth DP, Freebairn DM. 1996. APSIM: a novel software system for model development, model testing and simulation in agricultural systems research. *Agricultural Systems* 50(3):255–71
15. Ritchie JT. 1986. Model inputs. In *CERES-Maize: A Simulation Model of Maize Growth and Development*, eds. Jones CA, Kiniry JR. College Station, TX: Texas A&M University Press. pp. 37–48.
16. Ritchie JT. 1991. Specifications of the ideal model for predicting crop yields. In *Climatic risk in crop production: models and management for the semiarid tropics and subtropics*, eds. Muchow RC, Bellamy JA. Wallingford, UK: CAB International. pp. 97–122.
17. Jiang Z, Zhang J, Yang C, Rao Y, Liu S. 2016. Performance of classic multiple factor analysis and model fitting in crop modeling. *International Journal of Agricultural and Biological Engineering* 9:119–26
18. Silva RS, Kumar L, Shabani F, Picanço MC. 2018. An analysis of sensitivity of CLIMEX parameters in mapping species potential distribution and the broad-scale changes observed with minor variations in parameters values: an investigation using open-field *Solanum lycopersicum* and *Neoleucinodes elegantalis* as an example. *Theoretical and Applied Climatology* 132:135–44
19. Robertson MJ, Carberry PS, Huth NI, Turpin JE, Probert ME, et al. 2002. Simulation of growth and development of diverse legume species in APSIM. *Australian Journal of Agricultural Research* 53:429–46
20. Bhatia VS, Singh P, Wani SP, Chauhan GS, Rao AVRK, et al. 2008. Analysis of potential yields and yield gaps of rainfed soybean in India using CROPGRO-Soybean model. *Agricultural Forest Meteorology* 148:1252–65
21. Huang Y, Yu Y, Zhang W, Sun W, Liu S, et al. 2009. Agro-C: A biogeophysical model for simulating the carbon budget of agroecosystems. *Agricultural Forest Meteorology* 149:106–29
22. Wallach D, Goffinet B. 1987. Mean squared error of prediction in models for studying ecological and agronomic systems. *Biometrics* 43(3):561–73
23. Willmott C J. 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society* 3(11):1309–13
24. Brisson N, Ruget F, Gate P, Lorgeou J, Nicoullaud B, et al. 2002. STICS: a generic model for simulating crops and their water and nitrogen balances. II. Model validation for wheat and maize. *Agronomie* 22:69–92
25. Chauhan Y, Douglas C, Rachaputi R, Agius P, Martin W, et al. 2010. Physiology of mungbean and development of the mungbean crop model. In *Proceedings of the 1st Australian Summer Grains Conference, Gold Coas, Australia, 21–24 June 2010*. Collingwood, Australia: CSIRO.
26. Hammer GL, van Oosterom E, McLean G, Chapman SC, Broad I, et al. 2010. Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *Journal of Experimental Botany* 61(8):2185–202