Field Crops Research 119 (2010) 161–174

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Simulation of soybean growth and yield in near-optimal growth conditions

T.D. Setiyono[a], K.G. Cassman[a,*], J.E. Specht[a], A. Dobermann[b], A. Weiss[c], H. Yang[d], S.P. Conley[e], A.P. Robinson[f], P. Pedersen[g,1], J.L. De Bruin[h]

[a] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, P.O. Box 830915, Lincoln, NE 68583-0915, USA
[b] International Rice Research Institute, DAPO Box 7777, Metro Manila, Philippines
[c] School of Natural Resources, University of Nebraska-Lincoln, P.O. Box 830728, Lincoln, NE 68583-0728, USA
[d] Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, MO 6316, USA
[e] Department of Agronomy, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706, USA
[f] Department of Agronomy, Purdue University, 915 West State St., West Lafayette, IN 47907-2054, USA
[g] Department of Agronomy, Iowa State University, 2104 Agronomy Hall, Ames, IA 50011, USA
[h] Pioneer Hi-Bred International, Inc., 21888 N 950th Rd., Adair, IL 61411, USA

ARTICLE INFO

Article history:
Received 13 May 2010
Received in revised form 19 June 2010
Accepted 5 July 2010

Keywords:
Soybean
Glycine max
Biomass yield
Seed yield
Simulation model
Yield potential

ABSTRACT

SoySim is a new soybean (*Glycine max*, L. Merr) simulation model that combines existing approaches for the simulation of photosynthesis, biomass accumulation and partitioning with several new components: (i) flowering based on floral induction and post-induction processes, (ii) leaf area index based on logistic expansion and senescence functions, (iii) integration of canopy photosynthesis using a beta function, and (iv) yield simulation based on assimilate supply and seed number. Simulation of above ground dry matter (ADM) and seed yield by SoySim were validated against data from field studies at Lincoln (NE), Mead (NE), Whiting (IA), and West Lafayette (IN) that included 147 site-year-cultivar-planting date-plant-plant population combinations. In each of the four field studies, agronomic management other than planting date and plant population was optimized to achieve growth with minimal limitation from pests, nutrients, or other controllable factors. SoySim requires just two genotype-specific and two crop management-specific input parameters and yet provides reasonable accuracy in simulating growth and yield under optimum growth conditions across a wide range of sowing dates, plant population, and yield (2.5–6.4 Mg ha^{-1}) in the North-Central U.S. Corn Belt. Simulated seed yield had a RMSE of 0.46 Mg ha^{-1}. Few cultivar-specific parameter input requirements, lack of requirements for specification of key developmental stages, and mechanistic treatment of phenological development, canopy photosynthesis, and seed dry matter accumulation give several advantages to SoySim for use in research and for use as a decision-support tool to evaluate the impact of crop management options on yield potential in favorable environments.

1. Introduction

Yield potential is defined as the yield obtained when a crop is grown under field conditions with management practices that seek to eliminate growth reductions from nutrient deficiencies, insect pests, diseases, weeds, and moisture deficits or excess throughout crop growth—from planting to maturity (Evans, 1993). Under these favorable conditions, crop growth and yield are limited only by solar radiation and temperature. Soybean yields of 5.0–6.6 t ha^{-1} have been reported from a few replicated field studies in which researchers used crop and soil management practices to achieve yields near the yield potential ceiling, and yields at the high end of this range have been reported in sanctioned on-farm yield contests (Cooper, 2003; Specht et al., 1999). At issue is whether soybean growth simulation models have the ability to simulate soybean yields when the crop is grown in a manner that allows full expression of its yield potential.

A process-based soybean model must simulate phenology, biomass partitioning among organs, and yield formation. Accuracy in simulating biomass and grain yields across a wide range of environments requires an understanding of the biophysical determinants of these ecophysiological processes and incorporating such knowledge in mathematical formulations that constitute the model. To ensure that a model is credible in simulating yields over the full environmental range of possible yields, it is important to evaluate model performance under high-yield conditions where yields approach the yield potential ceiling, as well as in more stressful environments that produce lower yield levels.

Yet, relatively few publications documenting high-yield soybean environments (defined as yield at >4.5 Mg ha^{-1} range) can

* Corresponding author. Tel.: +1 402 472 5554; fax: +1 402 472 7904.
E-mail address: kcassman1@unl.edu (K.G. Cassman).
[1] Current address: Syngenta Crop Protection, 2415 Clayton Drive, Ames, IA 50010, USA.

0378-4290/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2010.07.007

Table 1
Description of datasets used for SoySim model validation.

Location[a]	Latitude, longitude, elevation	Tillage, soil series (family)	Years	Cultivars[b] (MG)	Sowing dates (month/day)	Plant population densities (plants m^{-2})	Seed yields (13% m.c.) (Mg ha^{-1})
Lincoln, NE (Setiyono et al., 2007)	40°50′10″N, 96°39′43″W, 357 m a.s.l.	Conventional tillage, Kennebec silt loam (fine-silty, mixed, superactive, mesic, Cumulic Hapluddols)	1999 through 2007	NE3001 (3.0) P93M11 (3.1) NE3201 (3.1) DKB 31-52 (3.1) L1067RR (3.1) NEX8903 (3.1) K323RR (3.2) P93B36 (3.3) KAUP335 (3.3) U98-307917 (3.4) AG3401 (3.4) S3632-4 (3.4) P93B47 (3.4) U98-307162 (3.6) U98-311442 (3.9)	4/26 to 6/14	21–45	2.45–5.90
Mead, NE (Verma et al., 2005)	41°9′53″N, 96°28′12″W, 351 m a.s.l.	No-till, Tomek silt loam (Fine, smectitic, mesic, Pachic Argiudolls)	2002 through 2006	P93B09 (3.0) P93M11 (3.1)	5/12 to 6/02	30–33	3.86–4.00
W. Lafayette, IN (Robinson et al., 2009)	40°28′7″N, 86°59′31″W, 190 m a.s.l	No-till, Drummer (Fine-silty, mixed, mesic, Typic Hapaquoll)	2006 through 2007	P92M61 (2.6) BECK321(3.2) BECK367 (3.6)	3/27 to 6/07	20–44	3.11–4.66
Whiting, IA (De Bruin and Pedersen, 2008)	42°8′1″N, 96°9′0″W, 323 m a.s.l	Conventional tillage, Salix silty clay loam (Fine-silty, mixed, mesic, Typic Hapludolls)	2004 through 2006	P91M90 (1.9) IA2068 (2.6) AG2801RR (2.8) PB291N (2.9)	4/27 to 6/06	13–56	3.72–6.39

[a] Experimental details not listed in table are provided in the cited publications.

[b] The first letter(s) of the cultivar brand name denote company/source that provided the soybean seed: NE (Univ. of Nebraska, Lincoln, NE), P (Pioneer Hi-Bred Intl. Inc., Johnston, IA), DKB (Dekalb, Monsanto Seed Co., Cortland, IL), L (Latham Hi-Tech Seeds, Alexander, IA), K (Kruger Seed Co., Dike, IA), KAUP (Kaup Seed, West Point, NE), U (Univ. of Nebraska, Lincoln, NE), AG (Asgrow, Monsanto Seed Co., Cortland, IL), S (Stine Seed Co., Adel, IA), BECK (Beck's Hybrids, Atlanta, IN), IA (Iowa State University, Ames, IA).

be found in the scientific literature (Salvagiotti et al., 2008). Given the limited availability of data, it is not surprisingly that the most widely used soybean models have not been validated against field observations from high-yield environments. Models such as Sinclair-Soybean (Sinclair, 1986) and CROPGRO-Soybean (Boote et al., 1998a) have, of course, been used to evaluate soybean production in many lower yield environments (1–3 Mg ha^{-1}), involving now-obsolete soybean cultivars (i.e. Williams and Bragg, released in the 1980s) grown at a limited number of sites, cultivars, and years (Boote et al., 1997; Carbone et al., 2003; Irmak et al., 2000; Jagtap and Jones, 2002; Muchow and Sinclair, 1986; Wang et al., 2003). And though Spaeth et al. (1987) did use high yield environments in their evaluation of Sinclair-Soybean, their experiments involved crop rotations, management practices and cultivars not typical of those widely used in the U.S. Corn Belt.

The U.S. Corn Belt region is a particularly important soybean-producing region because it accounts for about 75% of U.S. soybean production (2007–2009 National Agricultural Statistics Service, USDA, http://www.nass.usda.gov) and 30% of global production (2007 FAOSTAT, FAO, Rome, Italy, http://faostat.fao.org). Pedersen et al. (2004) used CROPGRO-Soybean to simulate the effect of environment and management on soybean yield under irrigated systems in north-central U.S. Corn Belt, and the authors noted the relatively poor agreement between observed and simulated yields (Fig. 3, Pedersen et al., 2004). Moreover, CROPGRO-Soybean requires a large number of cultivar-specific inputs to initiate the model, and gaining access to these parameters is difficult due to difficulties in measuring them. For example, while most seed companies provide information about the maturity group for each soybean cultivar they sell, they do not routinely provide, and may in fact not even have, information on cultivar-specific parameters required for CROPGRO-Soybean. Moreover, soybean cultivar turn-over in the North Central USA is rapid, which can render a cultivar-specific coefficient only transiently useful for model use. Lack of easy access to input parameters required by a crop model is a significant problem in that it limits the usefulness of that model in decision support for crop producers, industry professionals, and researchers.

Seed dry matter accumulation (hereafter referred to as seed growth) is particularly crucial for accurate simulation in high yield environments. This domain of relevance – a term used by Sinclair and Seligman (2000) in their discussion of new crop models – and the need to have a soybean model that more accurately simulate seed growth, motivated development of the SoySim model described in the present paper. When soybeans are grown with minimal limitation from biotic and abiotic stresses, we hypothesized that seed growth is no longer source-limited, but instead becomes sink-limited. Seed growth processes, however, are dynamic and influenced by reproductive duration and environmental factors such as solar radiation and temperature which, in turn, are influenced by sowing date (Bastidas et al., 2008). Existing models rely on a fixed rate of increase in harvest index during seed-filling (Sinclair-Soybean), a fixed partitioning coefficient (WOFOST, van Diepen et al., 1989), or a temperature-sensitive seed growth rate (CROPGRO-Soybean). At issue, then, is whether the approaches used in these three models are sufficiently robust for accurate prediction of soybean yields in high yield systems across a wide range of production environments in the North Central USA.

And while Sinclair-Soybean and WOFOST do not require a large number of cultivar-specific input parameters, *vis-a-vis* CROPGRO-Soybean, they do require specification of initiation of flowering and physiological maturity for the cultivar being simulated. Hence, neither model predicts phenological stages *ex ante*, which means in-season decision-support is not possible with either model. For example, some fungicides are the most effective when applied at the R3 stage (beginning pod stage, Fehr and Caviness, 1977), but not so effective earlier or later (Blaine et al., 2005; Trybom and Jeschke, 2008). Ability to project the calendar date of occurrence of this stage, including in-season continuous updating of that date using a combination of real-time and historical weather data would greatly facilitate the timely scheduling of labor and equipment for this field operation.

Given the lack of soybean model validation in high-yield environments, and the motivation to create a soybean model requiring only easily accessible cultivar-specific parameters, yet still providing in-season projections of soybean crop phenology based on real-time and historical weather data, our objectives were to (1) develop and validate a new soybean model – SoySim – designed with new formulations for key processes such as phenology, LAI, and seed growth while minimizing the need for cultivar-specific parameters, (2) compare the new model with the existing soybean models of CROPGRO-Soybean (representing a sophisticated crop model with extensive requirements for genotype-specific inputs), Sinclair-Soybean (representing a simpler model with few cultivar-specific inputs), and WOFOST (a generic crop model). As previously noted, the domain of relevance, for which SoySim was designed to accurately simulate, is a high-yield soybean environment. Such environments are common in many locations across the U.S. Corn Belt in most years, and particularly so in irrigated production systems used in western areas. The emphasis in the present study was to determine if SoySim, or any existing soybean model, could reliably simulate growth and yield response with varied genotype × environment × management practices. We speculate that improved formulations incorporated into SoySim for phenology, LAI, and seed growth would improve the simulation of soybean performance in these high-yield environments, even with relatively few cultivar-specific input parameters.

2. Materials and methods

2.1. Field experiments and photosynthesis measurements

Development and validation of SoySim relied on high quality data generated from well-managed field experiments supervised by agronomists who were attempting to use crop and soil management practices to minimize abiotic and biotic stresses. Such data were obtained from field experiments in Lincoln, NE (40°50′10″N, 96°39′43″W, 357 m above sea level), Mead, NE (41°9′53″N, 96°28′12″W, 351 m above sea level), W. Lafayette, IN (40°28′7″N, 86°59′31″W, 190 m above sea level), and Whiting, IA (42°8′1″N, 96°9′0″W, 323 m above sea level). Soybean (*Glycine max* L. Merr) seed yield (at 13% moisture) in these experiments ranged from 2.4 to 6.4 Mg ha^{-1}, with the lower yields associated with late sowing dates (Table 1). Experimental sites included both conservation tillage and no-till residue management. The soybean cultivars ranged in maturity group (MG) from 1.9 to 3.9, which are typical cultivar MGs used across the U.S. Corn Belt. Plant population densities at emergence varied from 13 to 56 plants m^{-2} (Table 1). Additional details about management of these field experiments were described in Setiyono et al. (2007), Verma et al. (2005), Robinson et al. (2009), and De Bruin and Pedersen (2008).

Relationships between net photosynthesis and intercepted photosynthetically active radiation (PAR) were evaluated in the 2004 Lincoln field experiment with cultivar NE3001. Leaf photosynthesis was measured with a portable photosynthesis system LI6400 (LI-COR, Lincoln, NE) on the uppermost fully expanded trifoliolate leaf at 43 and 71 days after emergence. Measurements were made between 11:00 and 14:00. Mean relative humidity was 55% and 57% on 43 and 71 days after emergence, respectively. On each day, measurements were taken on four representative plants from a treatment with optimal management (Setiyono et al., 2007). Mean

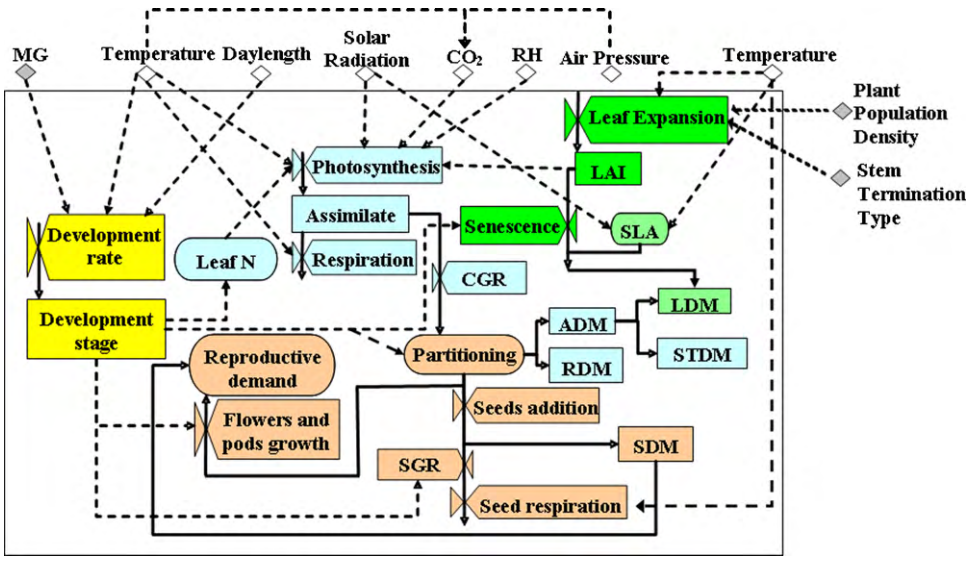

Fig. 1. Relational diagram of the SoySim model. The diagram follows the standard Forrester's system (Forrester, 1961), with slight modifications. The outer rectangular box encloses the system (soybean crop) with abiotic, genetic, and management factors that influence phenology and growth shown outside the box. MG = maturity group (number), plant population density (plants m^{-2}), stem termination type (indeterminate or semi-determinate), temperature (°C), daylength (h), solar radiation (MJ m^{-2} d^{-1}), RH = relative humidity (%), air pressure (mbar, estimated from site elevation), CGR = crop growth rate (kg ha^{-1} d^{-1}), LAI = leaf area index (m^2 m^{-2}), SLA = specific leaf area (cm^2 g^{-1}), ADM = total above ground dry matter (Mg ha^{-1}), RDM = root dry matter (Mg ha^{-1}), LDM = Leaf dry matter (Mg ha^{-1}), STDM = stem dry matter (Mg ha^{-1}), SGR = seed growth rate (kg ha^{-1} day^{-1}), SDM = seed dry matter (3% m.c., Mg ha^{-1}).

seed yield (at the standard 13% moisture content) was 4.9 Mg ha^{-1} in this treatment.

2.2. Model structure

The SoySim model includes a new approach for simulation of soybean dry matter production combined with previously developed phenology (Setiyono et al., 2007) and LAI (Setiyono et al., 2008) components. Phenology, LAI, and dry matter production components interact dynamically and each is influenced by weather variables (Fig. 1). Material flows in the relational diagram of the SoySim model in Fig. 1 are marked by a single assimilate deposit point driven by photosynthesis and several assimilate withdrawal points due to crop and seed maintenance respiration and leaf senescence, and utilization points governed by partitioning patterns and dry matter conversion efficiency. The model assumes that leaf and seed growth result from a dynamic interaction among the sink- and source-limited processes. The term 'limited' is used in the context of individual organ growth processes rather than at a whole plant level. Sink-limited growth describes growth associated with expansion of an organ as driven by cell division and expansion. When there is no water stress, these processes are influenced by temperature and phenology. Source-limited growth is associated with assimilate acquisition, loss, and partitioning, and is therefore governed by the amount of carbon resources available for allocation. The components of SoySim and their sources are described in Table 2. The current version simulates growth and yield in optimal conditions that allow expression of yield potential without abiotic and biotic stresses.

This paper highlights new formulations and approaches used in SoySim. A complete mathematical description of the model is provided in the SoySim user manual (available at http://soysim.unl.edu). New formulations in the SoySim model include: (1) explicit floral induction and evocation processes in simulation of flowering, and use of non-linear temperature and photoperiod functions for soybean developmental rates (Setiyono et al., 2007), (2) phenology- and temperature-driven leaf expansion and senescence processes with logistic functions, and a feedback effect of plant population density on maximum plant leaf area (Setiyono et al., 2008), (3) interacting effects of leaf N content, ambient [CO_2], PAR, air temperature, and RH on leaf photosynthesis rate, (4) a more accurate integration of photosynthesis from leaf to canopy level for a forbs-type leaf area density profile typical of the soybean canopy, and (5) seed dry matter accumulation driven by phenology and available assimilate during early reproductive stages. Descriptions of (3)–(5) of the above formulations novel to SoySim are provided in the following section.

The SoySim model is a Windows-based software and was written in Object Pascal language using the Delphi® 2007 for Win32 Integrated Development Environments (IDE).

2.3. Simulation of dry matter accumulation and partitioning

Above and below ground dry matter (*WDM*), aboveground dry matter (ADM), and root dry matter (RDM) in kg ha^{-1} were simulated using the framework of Supit and Van der Groot (2003):

$$\mathrm{WDM} = \mathrm{WDM}_{i-1} + (\mathrm{WGR})t \quad (1)$$

$$\mathrm{ADM} = \mathrm{ADM}_{i-1} + (1 - \pi_{\mathrm{RDM}})(\mathrm{WGR})t \quad (2)$$

$$\mathrm{RDM} = \mathrm{RDM}_{i-1} + \pi_{\mathrm{RDM}}(\mathrm{WGR})t \quad (3)$$

where WGR is growth rate of above and below ground dry matter (kg ha^{-1} d^{-1}), t is time interval (one day), π_{RDM} is the partitioning coefficient for root dry matter (kg kg^{-1}), and the subscript $i-1$ indicates the crop biomass (kg ha^{-1}) on the previous day. The root partitioning coefficient for root (π_{RDM}) was derived from a compilation of data obtained from multiple published reports (Fig. 2).

WGR is estimated by gross assimilation and maintenance respiration following Supit and Van der Groot (2003):

$$\mathrm{WGR} = Y_g(A - R_m) \quad (4)$$

where Y_g is a conversion efficiency factor for assimilate (kg kg^{-1}), A is the gross assimilation rate (kg ha^{-1} d^{-1}), and R_m is maintenance respiration rate (kg ha^{-1} d^{-1}). Y_g is defined as:

$$Y_g = \frac{1}{(\pi_{\mathrm{RDM}}(1 - \pi_{\mathrm{RDM}})/Y_{\mathrm{ADM}}) + (\pi_{\mathrm{RDM}}/Y_{\mathrm{RDM}})} \quad (5)$$

Table 2
Description and source (if not new) of major SoySim components.

Component	Source/Approach	Step	Driving variables
1. Vegetative developmental stages (V)	STICS[a], SOYDEV[b]	Daily	T_{mean}, P, MG, ST, R
2. Reproductive developmental stages (R)	SOYDEV[b]	Daily	T_{mean}, P, MG, ST, V
3. Leaf area index (LAI)	Logistic approach[c]	Daily	Phenology, T_{mean}
4.Dry matter accumulation			
a. Photosynthesis rate (P_s)	Farquhar[d], Harley[e], Yin[f], WOFOST[g], New[h]	Hourly	R_s, Phenology, RH, T, [CO_2], [O_2]
b. Respiration rate (R_s)	Farquhar, Harley[e], Yin[f], WOFOST[g]	Hourly	T, ADM, RDM, Phenology
d. Partitioning (below and above ground dry matter)	WOFOST[g], New[h,i]	Daily	Phenology
c. Crop growth rates (CGR)	WOFOST[g]	Daily	P_s, R_s, ADM
f. Yield formation			
i. Sink size determination	New[h]	Daily	Phenology, T_{mean}, CGR
ii. Seed growth rate (SGR)	New[h]	Hourly	Phenology, T_{mean}
5. Weather data input (conversion of daily to hourly weather data)			
i. Temperature (T)	WAVE[j]	Daily	T_{max}, T_{min}, P
ii. Solar radiation (R_s)	Michalsky[k], KASM-KC[l]	Daily	B, ATMTR, P
iii. Relative humidity (RH)	New[h]	Daily	Daily RH, T_{max}, T

MG = maturity group (number). ST = stem termination type. ATMTR = atmospheric transmission. B = solar elevation angle (°), P = daylength (h), T_{max} = maximum daily temperature (°C), T_{min} = minimum daily temperature (°C), T_{mean} = average daily temperature (°C), [CO_2] = ambient CO_2 concentration (375 μl l^{-1}), [O_2] = ambient CO_2 concentration (210 ml l^{-1}). ADM = above ground dry matter (Mg ha^{-1}). RDM = below ground (root) dry matter (Mg ha^{-1}).

[a] Brisson et al. (2003).
[b] Setiyono et al. (2007).
[c] Setiyono et al. (2008).
[d] Farquhar et al. (1980).
[e] Harley et al. (1985).
[f] Yin and van Laar (2005).
[g] Boogard et al. (1998).
[h] Setiyono et al. (this study).
[i] Mayaki et al. (1976), Cassman et al. (1980), Roder et al. (1989), Watt and Evans (2003), and Cheng et al. (2003).
[j] de Wit et al. (1978).
[k] Michalsky (1988).
[l] Badescu (1997).

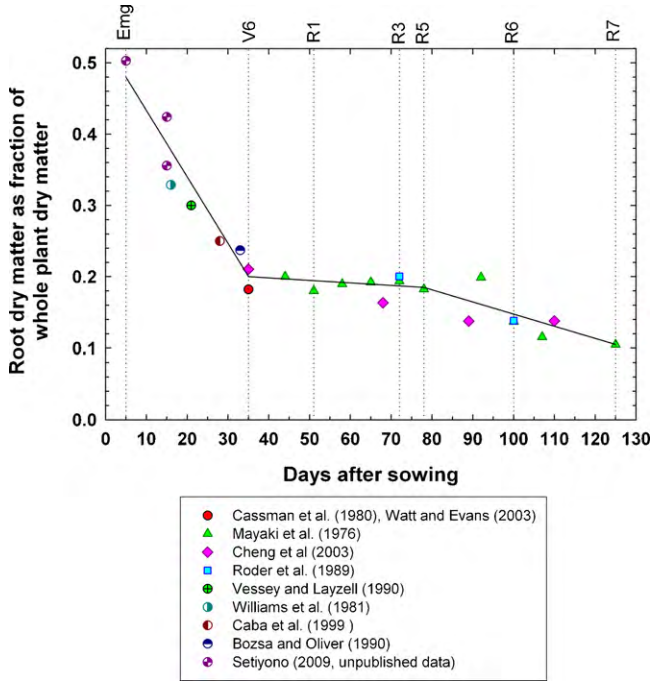

Fig. 2. Soybean root dry matter fraction of whole plant dry matter during the growing season based on Setiyono (2009, unpublished data, field experiment) and published data from experiments conducted in the field (Mayaki et al., 1976; Roder et al., 1989), greenhouse (Bozsa and Oliver, 1990; Cassman et al., 1980; Cheng et al., 2003; Watt and Evans, 2003), and growth chamber (Caba et al., 1999; Vessey and Layzell, 1990; Williams et al., 1981). Above-ground vegetative (V) and reproductive (R) developmental stages (Fehr and Caviness, 1977) are shown at top of the graph. The solid line is a composite mean of all reports, and represents the root dry matter fraction values used in SoySim.

where Y_{ADM} and Y_{RDM} are conversion efficiency factors for assimilate to ADM and RDM, respectively (kg kg^{-1}), which were estimated using values from WOFOST and the calibration dataset (Section 2.4). Eq. (5) is a simplified version of the function described in Supit and Van der Groot (2003) in that Eq. (5) does not account for biomass partitioning among structures such as stems, leaves, and storage organs. R_m was sensitive to temperature as follows:

$$R_m = R_{mr} Q_{10}^{(T-T_r)/10} \quad (6)$$

where R_{mr} is maintenance respiration (kg ha^{-1} d^{-1}) at a reference temperature (T_r) of 25 °C, Q_{10} is the fractional increase or decrease in respiration rate per 10 °C change in temperature, and T is temperature (°C). R_{mr} was simulated as a function of biomass and maintenance respiration coefficients for ADM and RDM:

$$R_{mr} = r_{ADM}\mathrm{ADM} + r_{RDM}\mathrm{RDM} \quad (7)$$

where r_{ADM} and r_{RDM} are maintenance respiration coefficients (kg kg^{-1} d^{-1}) for ADM and RDM, respectively. These coefficients were also estimated using values from WOFOST and the calibration dataset (Section 2.4).

Photosynthesis converts six moles of CO_2 into one mole of carbohydrate:

$$6CO_2 + 6H_2O \xrightarrow{\text{light}} C_6H_{12}O_6 + 6O_2 \quad (8)$$

thus the gross assimilation rate is calculated based on the carbon exchange rate as:

$$A = P_n \left(\frac{180}{264}\right) \quad (9)$$

where P_n is daily canopy photosynthesis rate (kg CO_2 ha^{-1} d^{-1}) and the term 180/264 represents the molecular weight ratio of one mole of carbohydrate to 6 moles of CO_2. The conversion term is commonly simplified to 30/44 kg carbohydrate kg^{-1} CO_2.

P_n is calculated by integrating the hourly net canopy photosynthesis (P_c) rate over the day. P_c is simulated by integration of individual leaf photosynthesis rates over the whole canopy using a multi-layer approach that distinguishes between sunlit and shaded leaves, and separation of PAR into direct and diffuse components (Goudriaan, 1986; Kropff and van Laar, 1993; Spitters et al., 1986; Supit and Van der Groot, 2003). Sunlit leaves receive both direct and diffuse components while shaded leaves receive only diffuse radiation (Supit and Van der Groot, 2003). Photosynthesis rate for sunlit and shaded leaves is simulated as function of absorbed PAR, ambient CO_2 concentration ($[CO_2]$), temperature, relative humidity (RH) and leaf N ($g\,m^{-2}$) using a modified C_3 plants photosynthesis model (Farquhar et al., 1980). Relationships between net photosynthesis and intercepted PAR or ambient $[CO_2]$, derived from measurements taken in this study (as mentioned earlier), were used in these functions, whereas the relationship between light-saturated photosynthesis and temperature was taken from Harley et al. (1985). Optimal leaf N (not limited by soil nutrient availability or symbiotic N fixation in root nodules) at different phenological stages was estimated from the data of Boote et al. (1998b), and the relationship between leaf photosynthetic rate and leaf [N] was taken from Sinclair and Horie (1989).

Seed dry matter is simulated by estimating the number of seeds and the mean individual seed growth rate:

$$\mathrm{SDM} = SW\,Sct\,0.01 \tag{10}$$

where SDM is seed dry matter ($kg\,ha^{-1}$), SW is mean individual seed weight (mg $seed^{-1}$), Sct is number of seeds ($seeds\,m^{-2}$) and the term 0.01 represents the unit conversion factor. Seeds still contain about 3% residual moisture when dried to constant weight in an oven at 70 °C, so for simulation purposes, seed yield at the standard 13% moisture content was calculated as: seed yield = (97/87)SDM.

The number of seeds per square meter is simulated as a function of assimilate availability, assimilate demand, and a partitioning coefficient using a modified Charles-Edwards model (Charles-Edwards et al., 1986):

$$Sct = Sct_{i-1} + (\Delta Sct)\mathrm{t} \tag{11}$$

$$\Delta Sct = \frac{\gamma_S \nabla F - R_{ms}}{a_{sct} Ad_s} \tag{12}$$

where ΔSct is the rate of change in the number of seeds ($seeds\,m^{-2}\,d^{-1}$), t is time interval (one day), γ_S is the partitioning coefficient for seeds (kg seed kg^{-1} aboveground biomass), ∇F is the assimilate available for growth ($kg\,ha^{-1}\,d^{-1}$). R_{ms} is the seed maintenance respiration ($kg\,ha^{-1}\,d^{-1}$), Ad_S is the assimilate demands required for continuing growth of an individual seed ($kg\,ha^{-1}$ $seed^{-1}\,d^{-1}$), a_{sct} is a proportional constant ($m^2\,d^{-1}$). Above-ground growth rate is used as an estimate ∇F (Charles-Edwards et al., 1986; Egli and Zhen-wen, 1991), and γ_S was simulated using a Gaussian function of developmental stage:

$$\gamma_S = a_{\gamma s} \exp\left(\frac{-(b_{\gamma s} - x_d)^2}{2(c_{\gamma s})^2}\right) \tag{13}$$

where x_d is a developmental progress variable for reproductive growth stage, $a_{\gamma s}$, $b_{\gamma s}$, $c_{\gamma s}$ are Gaussian coefficients for γ_S. Variable x_d is calculated as the cumulative rate from R1 to R7 based on temperature and daylength (value of zero at R1 and 1.0 at R7, Setiyono et al., 2007). From R7 to R8, the cumulative rate is continued to a value of 2.0 at R8. Ad_S is estimated from the reproductive growth rate, which is the sum of growth rates for seed and maternal tissue of the reproductive organs (flowers, pod walls, and seed coats):

$$Ad_S = \mathrm{MGR} + \mathrm{SGR} \tag{14}$$

where MGR is maternal tissue (flowers, pod walls, and seed coats) growth rate ($kg\,ha^{-1}\,d^{-1}$) and SGR is seed growth rate ($kg\,ha^{-1}\,d^{-1}$). The inclusion of maternal tissue for assimilate demand is based on Sheldrake's (1979) hypothesis that pods require a small but critical amount of assimilate during initial pod growth (during which the change in seed weight is negligible), followed by a much greater assimilate demand for seed growth. As a result, the availability of assimilate to support both sinks (maternal tissue and seeds) determines the daily number of seeds in the cohort that survive to contribute to the final seed number. SGR in Eq. (14) is calculated as the difference between the current and previous day SDM, while MGR is simulated using the first derivative of a logistic function and developmental stages:

$$\mathrm{MGR} = \frac{(a_{mgr}/b_{mgr})\exp(-(c_{mgr} - x_d)/b_{mgr})}{(1 + \exp(-(c_{mgr} - x_d)/b_{mgr}))^2} \tag{15}$$

where x_d is developmental progress from R1 to R7 and R8 (also used in Eq. (13)), and a_{mgr}, b_{mgr}, c_{mgr} are logistic function coefficients for MGR. These coefficients and those in Eqs. (17) and (18), were estimated from the calibration dataset (Section 2.4).

Individual seed weight, SW in Eq. (10), is estimated by the relationship:

$$SW = SW_{i-1} + \left(\Delta SW - \left(\frac{R_{ms}10^2}{Sct}\right)\right)\mathrm{t} \tag{16}$$

where ΔSW is gross rate of change in mean individual seed weight ($mg\,seed^{-1}\,d^{-1}$), the term i−1 represents the value of SW from the previous day, R_{ms} is seed maintenance respiration ($kg\,ha^{-1}$), Sct is number of seeds ($seeds\,m^{-2}$), the term 10^2 represents the unit correction factor ($[mg\,m^{-2}][ha\,kg^{-1}]$), t is time interval (one day). ΔSW is simulated using the first derivative of a logistic function and developmental stages:

$$\Delta SW = \frac{(a_{\Delta SW}/b_{\Delta SW})\exp(-(c_{\Delta SW} - x_d)/b_{\Delta SW}))}{\left(1 + \exp(-(c_{\Delta SW} - x_d)/b_{\Delta SW})\right)^2} \tag{17}$$

where x_d is developmental progress from R1 to R7 and R8 (also used in Eqs. (13) and (15)), $a_{\Delta SW}$, $b_{\Delta SW}$, and $c_{\Delta SW}$ are logistic function coefficients for ΔSW.

R_{ms} in Eqs. (12) and (16) is simulated as a function of temperature, seed maintenance respiration coefficient, and SDM as follows:

$$R_{ms} = (R_{msCref} Q10_{Rms}^{(T - T_{refrms})/10})\mathrm{SDM} \tag{18}$$

where R_{msCref} is seed maintenance respiration coefficient at reference temperature ($kg\,kg^{-1}\,d^{-1}$), $Q10_{rms}$ is fractional change in respiration rate with a 10 °C increase or decrease in temperature, T_{refrms} is reference temperature for seed maintenance respiration (25 °C), SDM is seed dry matter ($kg\,ha^{-1}$).

2.4. Model evaluation

The SoySim model was first calibrated against data from field experiments conducted in 2004 and 2005 at Lincoln, NE and in 2002 at Mead, NE. The calibration involved parameters governing relationships between maturity group and optimum developmental rates (Setiyono et al., 2007), plant population density and leaf expansion rates (Setiyono et al., 2008), and key biomass accumulation parameters that included maintenance respiration coefficients, conversion efficiency factor for assimilate, and partitioning coefficient for seed dry matter (this study). The revised final SoySim model was validated against data collected from field sites located in Lincoln NE (1999 through 2007), Mead NE (2002–2006), W. Lafayette IN (2006 and 2007), and Whiting IA (2004–2006) (Table 1).

For comparative purposes, soybean growth and yield at the validation sites were also simulated using CROPGRO-Soybean in the DSSAT version 4.0.2.0 software (Jones et al., 2003), Sinclair-Soybean (Sinclair, 1986; source code is available upon request), and WOFOST

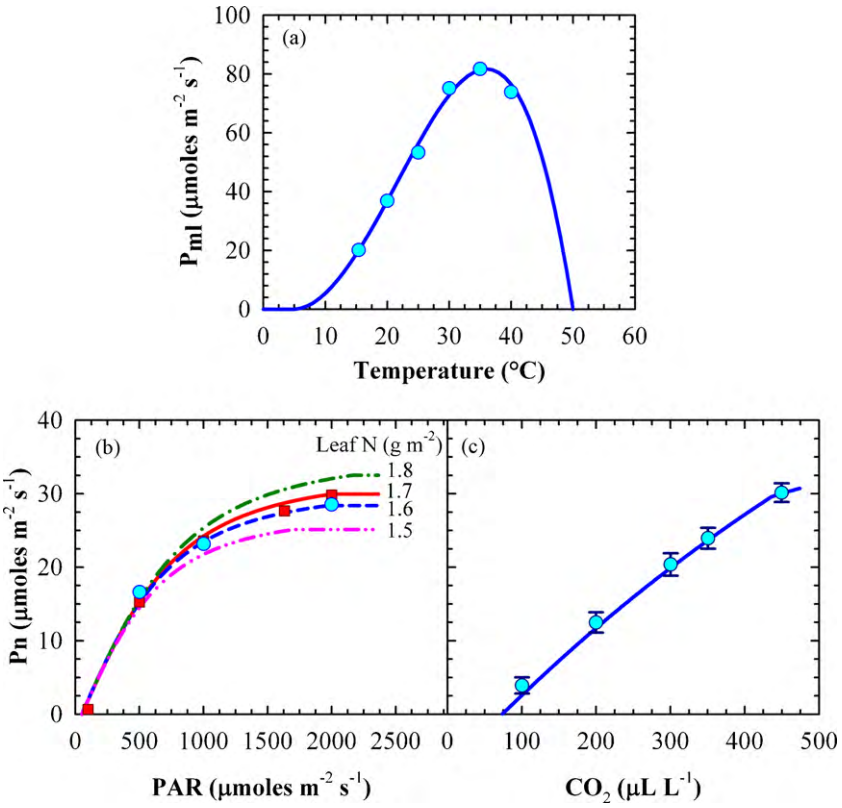

Fig. 3. Relationships governing photosynthesis rate (leaf-area basis) relative to three environmental variables as used in SoySim. Simulated responses are shown as lines in each of the three panels, whereas symbols represent observed published data or actual field measurements. Panels: (a) published observed data from Harley et al. (1985) and simulated light saturated photosynthesis (P_{ml}) in relation to temperature, (b) measured values from the current study and simulated net photosynthesis (P_n) in relation to intercepted PAR at four levels of leaf N, and (c) measured values from the current study versus ambient [CO_2] and simulated P_n versus [CO_2]. Measurements in (b) and (c) were taken at 43 (●, 54.9% RH) and 71 days after emergence (■, 57.4% RH) in Lincoln, NE in 2004. Standard errors are shown for each observed value, although some are too small to see relative to symbol size.

version 7.1 (Boogard et al., 1998). To ensure an impartial comparison, cultivar-specific inputs needed for the three models were first estimated using the same calibration dataset used for SoySim (Lincoln, NE in 2004 and 2005 and Mead, NE in 2002). During calibration of the three models, input parameters were optimized to allow each model to accurately simulate of phenology, LAI, ADM, and final seed yield in that calibration dataset, following the same method as that used in Hunt and Boote (1998).

Root mean square error (RMSE) and mean error (ME) were calculated for simulated values from all models following the methods given in Janssen and Heuberger (1995):

$$\mathrm{RMSE} = \sqrt{\frac{\sum (s_i - o_i)^2}{n}} \quad (19)$$

$$\mathrm{ME} = \frac{\sum (s_i - o_i)}{n} \quad (20)$$

where s = simulated data for the ith site-year-cultivar-planting date-plant population combination (experimental unit), o = observed data on the ith experimental unit, n = number of pairs of simulated and observed data.

3. Results

3.1. Model description and calibration

Model-based estimation of leaf photosynthesis rate under optimum conditions must account for interacting influences of temperature, light intensity, ambient [CO_2], and leaf N. Although the SoySim model utilizes the relationship between *maximum* pho-

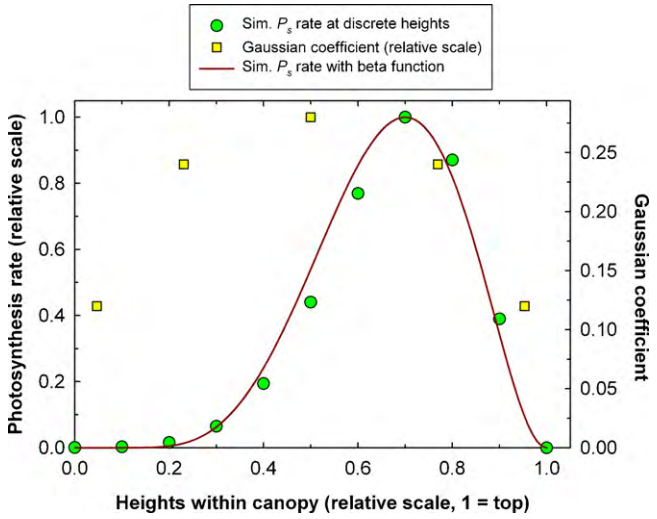

Fig. 4. Simulation of soybean photosynthesis rate (ground-area-basis) as function of heights within a soybean canopy. Symbols are P_s rates simulated with SoySim at discrete elevations within the canopy (●) and Gaussian coefficients (Goudriaan, 1986; Kropff and van Laar, 1993) for multilayer integration of P_s rate (■). Line is simulated P_s rates with beta function (Yin et al., 1995) with cardinal elevations of 0.1, 0.7, and 1.0 (r_{min}, r_{opt}, and r_{max}, respectively). The integrated P_s rate with this beta function was 0.36 (on a relative scale). With a maximum P_s rate (at r_{opt}) of 600 g CO_2 ha^{-1} h^{-1}, the integration with this beta function resulted in 216 g CO_2 ha^{-1} h^{-1} of canopy photosynthesis.

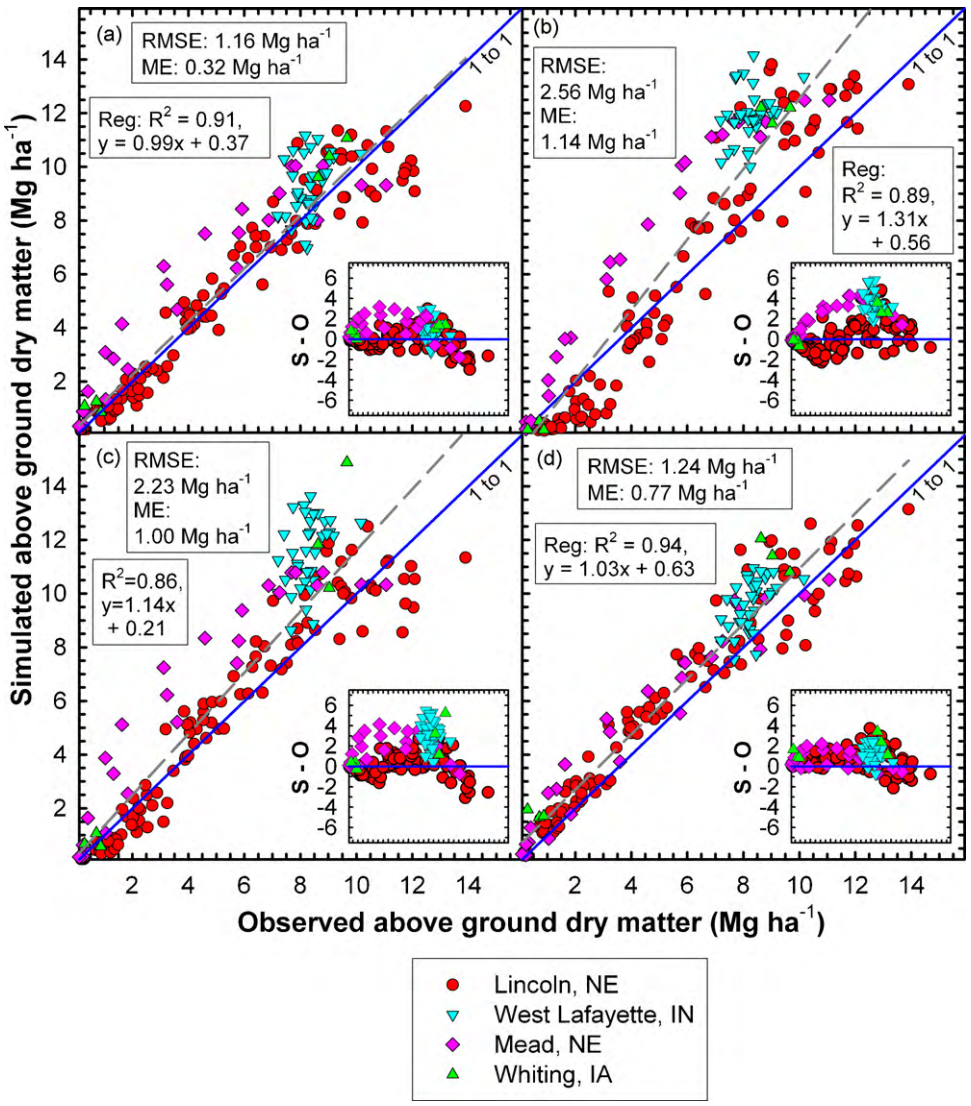

Fig. 5. Soybean above-ground dry matter as observed (O) and simulated (S) by the (a) CROPGRO-Soybean, (b) Sinclair-Soybean, (c) WOFOST, and (d) SoySim models. RMSE = root mean square error, ME = mean error, Reg = regression line (dashed line). Inserts are plots of residual (S–O) versus observed above ground dry matter. Observed values represent measurements taken at different growth stages up to physiological maturity. Data shown were of progressive crop growth.

tosynthetic rate at *optimal* light intensity as defined by Harley et al. (1985, Fig. 3a), actual data from our field experiments in Lincoln NE were used to establish the response of photosynthesis to light intensity at optimal temperature (Fig. 3b), and response of photosynthesis to [CO_2] levels (Fig. 3c). Cardinal temperatures for the photosynthesis response curve were 5, 36, and 50 °C for T_{min}, T_{opt}, and T_{max}, respectively. Small differences in leaf N (g m^{-2}) may explain some of the variability in observed data that were used to define the light response curve as evident by the simulated P_n versus intercepted *PAR* in Fig. 3b. In contrast to the response to intercepted light, the response of P_n to [CO_2] did not show saturation up to highest CO_2 treatment of 400 μL L^{-1} (Fig. 3c).

The standard Gaussian approach of integrating photosynthesis from individual leaves to whole canopy (Goudriaan, 1986; Kropff and van Laar, 1993) was not compatible with the photosynthesis rate distribution along the vertical canopy axis of a soybean plant community. The Gaussian approach gives greater weight to photosynthesis in the middle of canopy (Fig. 4), whereas leaf area distribution in soybean is more heavily weighted towards the upper canopy (field observations). Beta functions (Wang and Engel, 1998; Yin et al., 1995) describe a canopy photosynthesis profile that is more similar to soybean leaf area distribution compared to Gaussian coefficients (Fig. 4). Thus, a Beta function was used in SoySim to integrate P_s rate along the vertical canopy axis.

SoySim calibrations included phenology parameters governing the relationship between maturity groups and maximum developmental rates for key developmental stages (Setiyono et al., 2007), and growth functions relating LAI parameters to stem termination type (indeterminate or semi-determinate) and plant population density (Setiyono et al., 2008). As a result, SoySim requires specification of only two cultivar-specific parameters: maturity group (MG) and stem termination growth habit (i.e. semi-determinate or indeterminate types, though the latter is the predominant cultivar type available in most North Central USA regions). MG information is typically embedded in the brand name of proprietary cultivars and certainly easily obtainable from company web sites and seed dealers. Other growth parameters, including those governing the canopy leaf density profile, photosynthesis, respiration, and partitioning, are assumed in SoySim to be non-cultivar specific for commonly used cultivars within the MG range of 1.9–3.9. The

Table 3
Comparative list of inputs and key-processes algorithms for simulation of soybean yield potential in the four crop models evaluated in this study.

Aspects	Models			
	CROPGRO-Soybean[a]	Sinclair-Soybean[b]	WOFOST[c]	SoySim[d]
Weather inputs	• Latitude (°), longitude (°) • Solar radiation (MJ m^{-2}) • T_{max} (°C) • T_{min} (°C)	• Solar radiation (MJ m^{-2}) • T_{max} (°C) • T_{min} (°C)	• Longitude (°), latitude (°), elevation (m) • Solar radiation (kJ m^{-2}) • T_{max} (°C) • T_{min} (°C)	• Latitude (°), longitude (°), elevation (m) • Solar radiation (MJ m^{-2}) • T_{max} (°C) • T_{min} (°C) • RH (%)
Cultivar-specific inputs				
Phenology	(1) SOW to EMG (td) (2) EMG to V1 (td) (3) Critical short day length (h) (4) Photoperiod sensitivity (h^{-1}) (5) V-stages (leaves (td)$^{-1}$) (6) EMG to R1 (ptd) (7) R1to R3 (ptd) (8) R1 to R5 (ptd) (9) R5 to R7 (ptd) (10) R1 to end-of-leaf (ptd) (11) Seed filling duration (ptd) (12) Pod filling duration (ptd)	(1) SOW to V1 (°Cd) (2) PLI coefficient ([°Cd]$^{-1}$) (3) SOW to end-of-leaf (d) (4) SOW to R5 (d)	(1) SOW to EMG (°Cd) (2) EMG to R1 (°Cd) (3) R1 to R7 (°Cd)	(1) Maturity group number (2) Type of stem termination
Leaf growth	(1) SLA (cm^2 g^{-1}) (2) Maximum leaf size (cm^2) (3) V-stage when leaf growth switch from sink to source driven (VSINK) (4) Coefficients (a, b, c, d, e) relating V-stage and leaf area per plant	Exponential function coefficients (a, b, and c) for plant leaf area and PLI relationship.	(1) LAI at EMG (ha ha^{-1}) (2) Maximum relative increase in LAI (ha ha^{-1} d^{-1}) (3) SLA (ha kg^{-1}) (4) DM conversion efficiency for leaves (kg kg^{-1}) (5) DM partitioning coefficient for leaves (kg kg^{-1})	Type of stem termination
Seed growth	(1) Maximum weight per seed (g) (2) Average seed number per pod (seeds pod^{-1}) (3) Maximum seed growth rate (kg ha^{-1} d^{-1})	(1) Initial HI (2) Rate of increase in HI (d^{-1}) (3) Seed nitrogen (g N g^{-1})	(1) DM conversion efficiency for seeds (kg kg^{-1}) (2) DM partitioning coefficient for seeds (kg kg^{-1})	Type of stem termination
Model algorithms				
Phenology	Key stages: (1) EMG, (2) R1 (3) R3, (4) R5, (5) end-of-leaf (6) R7, (7) R8. Approach: linear spline function with four cardinal values (T_{mean}) and inverse linear function with two coefficients (P).	Key stages: (1) V1, (2) end-of-leaf, (3) R5. Approach: thermal time with a T_{base}.	Key stages: (1) EMG, (2) R1, (3) R7. Approach: thermal time with T_{base} and T_{opt}.	Key stages: (1) EMG, (2) V-stages, (3) R1, (4) R3, (5) R5, (6) R7, (7) R8. Approach: non-linear functions of T_{mean} and P with distinct floral induction and evocation processes for flowering and dynamic interactions between V-stages and R-stages.
Leaf growth	Leaf growth is sink-driven prior to VSINK stage and source limited thereafter. Sink-driven leaf area expansion is an exponential function of cultivar coefficients and V-stages. Leaf area after VSINK depends on the partitioning of dry matter to leaf, specific leaf area of new leaf tissue, and total carbohydrates available for growth.	Leaf growth is driven by PLI and plant population density until end-of-leaf stage. Thereafter reduction in leaf area is driven by N remobilization from leaves as results of N demand from the growing seed.	Leaf growth is driven by minimum of either sink or source driven processes. Sink-limited leaf growth is based on the maximum rate of leaf area expansion (a cultivar specific parameter) and temperature using an exponential function. Source-limited leaf expansion is driven by SLA, partitioning coefficients of dry matter to leaf, and available dry matter for growth.	Leaf growth is driven by simultaneous leaf expansion and senescence processes at the whole plant and population level. The rates of leaf area expansion and senescence are simulated using logistic functions taking into account that high plant population density reduces the maximum leaf area per plant due to inter-plant competition effect.

Table 3 (*Continued*)

Aspects	Models			
	CROPGRO-Soybean[a]	Sinclair-Soybean[b]	WOFOST[c]	SoySim[d]
Crop growth	Crop growth is driven by photosynthesis, respiration, and detailed carbon and nitrogen utilization and allocation processes. Simulation of maintenance respiration considers the effect of crop dry weight and photosynthesis activity. Crop photosynthesis is simulated based on hourly leaf level and hedgerow photosynthesis.	Crop growth is driven by intercepted solar radiation and RUE (1.2 g biomass/MJ intercepted irradiance). Partitioning of ADM to stem and leaf is governed by the specific leaf N contents in each organ.	Crop growth is driven by photosynthesis and respiration. Leaf photosynthesis rate is calculated as a function of maximum photosynthesis, initial light use efficiency, and absorbed PAR, and is integrated into canopy level using Gaussian approach. Net assimilation is calculated based on gross photosynthesis, maintenance respiration (T_{mean}-driven), and growth respiration (phenology-driven).	Crop growth is driven by photosynthesis and respiration. Photosynthesis rate at the individual leaf level is influenced by solar radiation, air temperature, leaf-N dynamic regulated by crop phenology, RH, and [CO_2]. Leaf photosynthesis rate is integrated from leaf into canopy scale using beta function. Net assimilation is calculated based on gross photosynthesis, maintenance respiration (temperature-driven), and growth respiration (phenology-driven).
Seed growth	Seed growth is driven by crop C and N status, partitioning coefficients, and production cost of pods and seeds. Flower, pod, and seed addition, growth, and abortion are simulated. Number of pods is governed by air temperature and pod filling duration (cultivar- specific coefficient). Number of seeds is calculated based on number of seeds per pod (cultivar-specific input), while seed growth rate is simulated based on maximum seed growth rate (cultivar-specific input). Both number of seeds and seed growth rate are influenced by air temperature and *P*.	Seed growth is driven by ADM and HI increase rate (0.11 kg ha^{-1} d^{-1}). Seed growth removes nitrogen from leaves and stems. The reduction in leaf N results in leaf senescence, which in turn results in leaf abscission causing an additional loss in nitrogen and vegetative dry matter.	Seed growth is driven by crop ADM accumulation and partitioning coefficient to seeds. Partitioning coefficients, including for other organs such as stems, leaves, and roots are phenology-driven.	Seed growth is simulated based on number of seeds (driven by crop growth rate during early reproductive stage, Charles-Edward et al., 1986) and individual seed growth corrected for maintenance respiration.

Abbreviations: (1) Weather variables: *P*, daylength; RH, relative humidity; T_{mean}, daily average air temperature; T_{max}, daily maximum air temperature; T_{min}, daily minimum air temperature; (2) Model variables: DM, dry matter; HI, harvest index; ptd, photothermal day; RUE, crop radiation use efficiency; SLA, specific leaf area; T_{base}, base mean temperature for crop development; td, thermal day; ADM, above ground dry matter; T_{opt}, optimum average temperature for crop development; (3) Developmental stages: end-of-leaf, end of leaf development; EMG, emergence; PLI, plastocrhon index; R1, first flower; R3, first pod; R5, first seed; R7, physiological maturity; R8, harvest maturity; SOW, Sowing; V1, first leaf.

[a] Boote et al. (1998a).
[b] Sinclair (1986).
[c] Supit and Van der Groot (2003).
[d] Setiyono et al. (this paper).

overall RMSE for SoySim simulations of ADM and SDM with the calibration dataset was 0.59 and 0.09 Mg ha^{-1}, respectively.

3.2. Model validation

SoySim requires fewer cultivar-specific inputs compared to CROPGRO-Soybean, Sinclair-Soybean, and WOFOST as summarized in (Table 3). Despite use of just two cultivar-specific inputs, SoySim simulated ADM and seed yield with similar or better accuracy compared to the other models (Figs. 5–7).

SoySim simulation of ADM with the validation datasets had an RMSE of 1.24 Mg ha^{-1} versus the respective RMSE values of 1.16, 2.56, and 2.23 Mg ha^{-1} for simulations produced by CROPGRO-Soybean, Sinclair-Soybean, and WOFOST (Fig. 5). The positive mean error (ME), which is indicative of overestimation bias, was evident in the ADM simulations of all four models, and especially pronounced in the Sinclair-Soybean and WOFOST simulations. Distribution of differences between simulated and observed ADM throughout the range of observed ADM values (inset graphs in Fig. 5) had a standard deviation (s.d.) of 0.97 Mg ha^{-1} for SoySim, compared to s.d. values of 1.12, 2.01, and 1.75 Mg ha^{-1} for CROPGRO-Soybean, Sinclair-Soybean, and WOFOST, respectively.

The validation database in this study provided a wide range of seed yield (from 2.4 to 6.4 Mg ha^{-1}). Differences in observed seed yields were primarily due to different sowing-date treatments at some of the sites. Greater yields were associated with an earlier sowing date. Simulation of seed yield across the validation sites with CROPGRO-Soybean, Sinclair-Soybean, WOFOST, and SoySim resulted in RMSE values of 0.72, 0.91, 1.61, and 0.46 Mg ha^{-1}, respectively (Fig. 6). Again, positive ME values indicated that simulated seed yields were, on average, overestimated by all four models (inset graphs in Fig. 6). Linear regression of simulated seed yield on observed seed yield generated regression coefficients that were less than unity for all four models. However, the coefficient for SoySim was substantively closer to unity (i.e. 0.94) than those of the other three models (0.26, 0.12, and 0.26). The cross-over point between regression trend line and 1:1 line indicates that seed yield begins to be underestimated when observed yields exceed about 4.5 Mg ha^{-1} in the CROPGRO-Soybean simulations, and about 5.0 Mg ha^{-1} in the Sinclair-Soybean simulations. Seed yield simulated by WOFOST was generally overestimated at all observed yield levels. SoySim simulated the variation in observed yields reasonably well and with minimal bias across the range of measured values, as indicated by close agreement between the linear regression line and

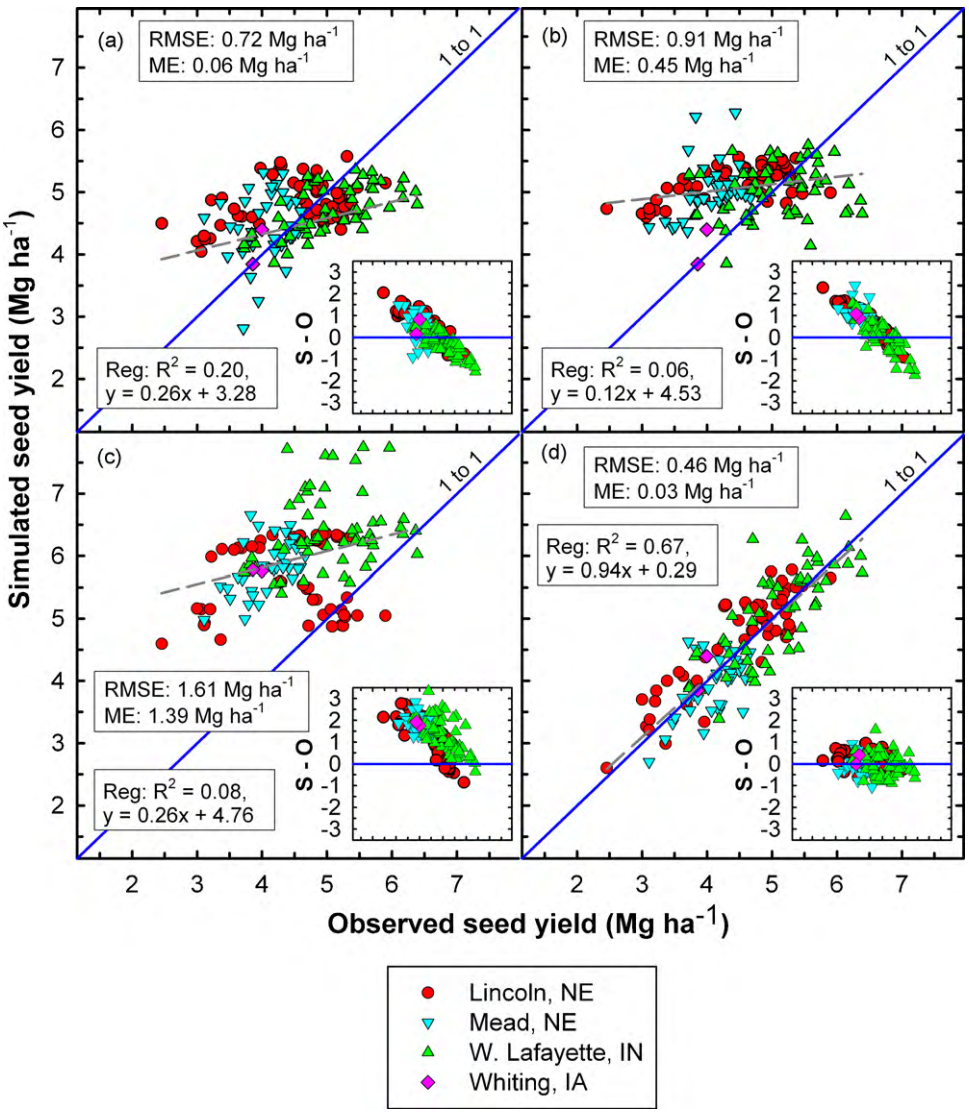

Fig. 6. Soybean seed yield (13% moisture content) as observed (O) and simulated (S) by the (a) CROPGRO-Soybean, (b) Sinclair-Soybean, (c) WOFOST, and (d) SoySim models. RMSE = root mean square error, ME = mean error, Reg = regression line (dashed line). Inserts are plots of residual (S–O) versus observed seed yield.

the 1:1 line (Fig. 6). Distribution of differences between observed and simulated seed yields throughout the range of observed seed yields (inset Fig. 6) showed a distinct declining trend at higher seed yields with CROPGRO-Soybean and WOFOST. The s.d. values of these differences were 0.72, 0.77, and 0.80 Mg ha^{-1}, respectively, for simulations with CROPGRO-Soybean, Sinclair-Soybean, and WOFOST. Observed versus simulated differences in seed yield with SoySim had considerably smaller s.d. (0.46 Mg ha^{-1}). Compared to the simulated yields generated by SoySim, those generated by the other models were clearly not as effective in predicting the observed yield values in this multiple-location validation data set.

The SoySim model also provided reasonable simulation of the ADM accumulation pattern (Fig. 7). In most cases, the simulated ADM pattern obtained with SoySim aligned more closely with the observed ADM pattern than did simulation patterns generated with the other models, especially late in the growing season. SoySim simulated late season decline in ADM between R7 and R8 similar to that of CROPOGR-Soybean and Sinclair-Soybean. A decline in observed ADM from R7 to R8 is consistent with senescence and abscission of leaves and petioles (i.e. lost dry matter) that occurs in a soybean crop during this phase.

4. Discussion

The SoySim model simulates canopy and seed growth with reasonable accuracy across a wide range of environments and yield potential in the U.S. Corn-Belt. For example, in the validation data set at locations in IA, IN, and NE, observed seed yields ranged from 2.45 to 6.39 Mg ha^{-1}, sowing dates were as early as late March to as late as mid June, plant population density varied from 13 to 56 m^{-2}, and cultivars differed in maturity from MG 1.9 to MG 3.9. SoySim achieves this robust simulation capability with just two cultivar-specific inputs and without the need for *a priori* specification of calendar dates for the key developmental stages. New formulations in SoySim provide for a more mechanistic treatment of phenological development, leaf expansion and senescence, leaf and canopy photosynthesis, and seed dry matter accumulation driven by phenology and availability of assimilates during early reproductive stage. SoySim's mechanistic treatment of these processes was expected to improve the simulation of seed yield and leaf area with fewer input parameter requirements relative to the other models evaluated in this study. For example, SoySim simulates LAI by simultaneously accounting for both leaf expansion and senescence processes at the whole plant and plant community

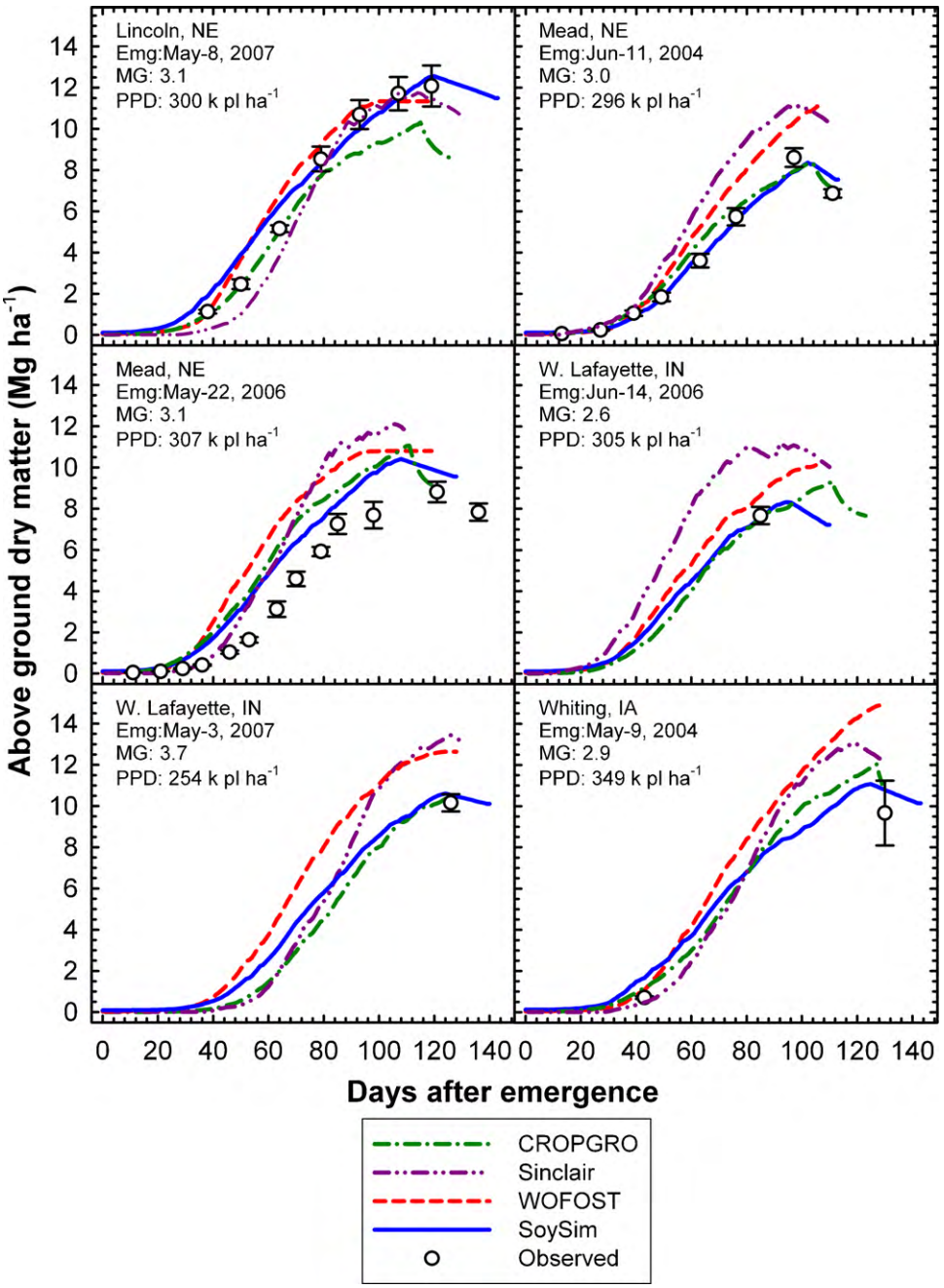

Fig. 7. Observed and simulated above ground dry matter (ADM) accumulation during the growing season in selected site × year × treatments combinations. Simulations were performed using the CROPGRO-Soybean, Sinclair-Soybean, WOFOST, and SoySim models. PPD = Plant population density (plants ha^{-1}).

level. Use of a logistic function for expansion and senescence allows the termination of leaf growth to be a model output rather than the required model input specification needed in CROPGRO-Soybean and Sinclair-Soybean (Table 3).

SoySim also avoids the need for difficult-to-measure (or -acquire) cultivar-specific parameters such as specific leaf area (SLA) required in WOFOST, and the elapsed time to when leaf growth switches from sink- to source-driven growth required in CROPGRO-Soybean (Table 3). In addition, SoySim does not require cultivar-specific parameters for seed growth that are required in CROPGRO-Soybean, which necessitates specification of pod-filling duration, number of seeds per pod, and maximum seed growth rate. While cultivars can differ in these parameters, those genotypic differences are not necessary invariant to environmental influences arising from location and management decisions, which makes difficult the acquisition of pertinent parameter values in the case of unplanned (i.e. on-the-fly) simulations. In contrast, robust simulation of seed yield by SoySim across a wide range of North Central USA environments was due in part to simulation of seed number response to environmental conditions as influenced by local weather and specific crop management. Seed number was measured in a just five (validation) locations, where measured seed yields ranged from 3.4 to 5.0 Mg ha^{-1} and seed number ranged from 2322 to 3512 ha^{-1}. There was a strong correlation between the two variables ($r = 0.913$, $n = 5$). The SoySim simulated number of seeds ranged from 2217 to 4379 seeds m^{-2} (RMSE = 625 seeds m^{-2}). De Bruin and Pedersen (2008) documented a strong positive correlation between soybean seed yield and seed number across years, sites, and cultivars in Iowa.

The seed yield component of SoySim is not based on empirical relationships between environmental factors and seed growth rates. Instead, seed growth is simulated as an integrated sink within the plant (Farrar, 1996) that, in turn, has a minimum growth requirement to sustain pod viability (Sheldrake, 1979), and a mechanistically determined final seed number (Charles-Edwards et al., 1986). This more fundamental approach captures the influence of environmental factors on seed growth in a dynamic fashion, thereby providing SoySim user with an opportunity to perform sensitivity analysis of pod development and seed abortion to identify (and thus seed genetic or agronomic mitigation of) the possible barriers to higher soybean yield potential (Cooper, 2003), or to evaluate the benefit of early sowing for optimizing number of main stem nodes for high yield (Bastidas et al., 2008). In SoySim, the simulation formulations for seed number determination are sensitive to crop growth rate during early reproductive stage (R1–R4)—a sensitivity documented by Egli and Zhen-wen (1991). Indeed, in the rice model ORYZA2000, accurate yield simulation relies on the number of spikelets, which is based on crop growth rate in the period from panicle initiation to first flowering (Bouman et al., 2001).

With access to historical weather data of adequate duration (e.g. 20–30 years) near the field site, output from the SoySim model allows the user to evaluate soybean yield potential at a given location with respect to any choice of planting date, North Central USA cultivar maturity group, and plant population. Because the model utilizes dynamic relationships for control of photosynthesis in relation to [CO_2], temperature, leaf nitrogen, solar radiation, and relative humidity (Farquhar et al., 1980; Harley et al., 1985; Sinclair and Horie, 1989; Yin and van Laar, 2005), and employs a multilayer integration of canopy photosynthesis (Kropff and van Laar, 1993; Supit and Van der Groot, 2003), it could be used to examine interactive effects of planting date, plant population and cultivar MG on yield potential in favorable environments, explore the impact of climate change on yield potential, and to assist breeders in identifying optimal canopy architecture and growth habit for higher yield potential.

5. Conclusions

Results from the present study document that the SoySim model met the criteria proposed by Sinclair and Seligman (2000) for publishing a new crop model. The new model has a well-defined domain of relevance (high yield environments), a mechanistic framework for simulation of crop growth and yield, and our analysis herein included an evaluation of scientific innovations embodied in a new model. In contrast to existing models, SoySim uses new mechanistic formulations to simulate phenology, canopy photosynthesis, and dry matter accumulation. Because SoySim requires just two readily available cultivar-specific inputs and no *a priori* specification of key developmental stages, the model has potential for a broad range of uses as a decision-support tool to improve soybean producer crop management and to support production-based agronomic and cultivar development research.

Acknowledgements

Funding for this research was provided by the Nebraska Soybean Board, the Fluid Fertilizer Foundation (FFF), the Foundation for Agronomic Research (FAR), the International Plant Nutrient Institute (IPNI), and the Agricultural Research Division, Institute of Agriculture and Natural Resources, University of Nebraska. The authors express their appreciation to Tim Arkebauer, David Scoby, Fernando Salvagiotti, Angela Bastisdas, Joe Mitchell, Travis Wegner, Darren Binder, and Tanvir Shah for their help with field operations and measurements, to John Linquist for creative suggestions on model development, and to Daniel Ginting and Haishun Yang for their help in computer programming. The authors also thank the reviewers for their careful and thorough examination of this paper and for their constructive comments and insights.

References

Badescu, V., 1997. Verification of some very simple clear and cloudy sky models to evaluate global solar irradiance. Solar Energy 61, 251–264.

Bastidas, A.M., Setiyono, T.D., Dobermann, A., Cassman, K.G., Elmore, R.W., Graef, G.L., 2008. Soybean sowing date: the vegetative, reproductive, and agronomic impacts. Crop Sci. 48, 727–740.

Blaine, A., Fulton, H., Graves, B., Henn, A., Ingram, D., McDaniel, T., Moore, W.F., Poston, D., Sciumbato, G., Spinks, B., Ward, B., Wardlaw, M., Zhang, L., 2005. Suggested Guidelines for Using Fungicides to Manage Soybean Rust. Mississippi State University Extension Service.

Boogard, H.L., van Diepen, C.A., Rötter, R.P., Cabrera, J.M.C.A., van Laar, H.H., 1998. User's guide for the WOFOST 7.1 crop growth simulation model and WOFOST Control Center 1.5. Technical Document 52. DLO Winand Staring Centre, Wageningen, The Netherlands.

Boote, K.J., Jones, J.W., Hoogenboom, G., 1998a. Simulation of crop growth: CROPGRO model. In: Peart, R.M., Curry, R.B. (Eds.), Agricultural Systems Modeling and Simulation. Marcel Dekker, Inc., New York, USA, pp. 651–692.

Boote, K.J., Jones, J.W., Hoogenboom, G., Pickering, N.B., 1998b. The CROPGRO model for grain legumes. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Norwell, MA, pp. 99–128.

Boote, K.J., Jones, J.W., Hoogenboom, G., Wilkerson, G.G., 1997. Evaluation of the CROPGRO-Soybean model over a wide range of experiments. In: Kropff, M.J., Teng, P.S., Aggarwal, P.K., Bouma, J., Bouman, B.A.M., van Laar, H.H. (Eds.), Applications of Systems Approaches at the Field Level. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 113–133.

Bouman, B.A.M., Kropff, M.J., Toung, T.P., Wopereis, M.C.S., ten Berge, H.F.M., van Laar, H.H., 2001. ORYZA2000: Modeling Lowland Rice. International Rice Research Institute, Los Banos, Philippines.

Bozsa, R.C., Oliver, L.R., 1990. Competitive mechanism of common cocklebur (*Xanthium strumarium*) and soybean (*Glycine max*) during seedling growth. Weed Sci. 38, 344–350.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussiere, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillere, J.P., Henault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model STICS. Eur. J. Agron. 18, 309–332.

Caba, J.M., Poveda, J.L., Gresshoff, P.M., Ligero, F., 1999. Differential sensitivity of nodulation to ethylene in soybean cv. Bragg and a supernodulating mutant. New Phytol. 142, 233–242.

Carbone, G.J., Mearns, L.O., Mavromatis, T., Sadler, E.J., Stooksbury, D., 2003. Evaluating CROPGRO-soybean performance for use in climate impact studies. Agron. J. 95, 537–544.

Cassman, K.G., Whitney, K.R., Stockinger, K.R., 1980. Root growth and dry matter distribution of soybean as affected by phosphorus stress, nodulation, and nitrogen source. Crop Sci. 73, 17–22.

Charles-Edwards, D.A., Doley, D., Rimmington, G.M., 1986. Modelling Plant Growth and Development. Academic Press, North Ryde, NSW.

Cheng, W., Johnson, D.W., Fu, S., 2003. Rhizosphere effects on decomposition: controls of plant species, phenology, and fertilization. Soil Sci. Soc. Am. J. 67, 1418–1427.

Cooper, R.L., 2003. A delayed flowering barrier to higher soybean yields. Field Crop Res. 82, 27–35.

De Bruin, J.L., Pedersen, P., 2008. Soybean seed yield response to planting date and seeding rate in the Upper Midwest. Agron. J. 100, 696–703.

de Wit, C.T., Goudriaan, J., Laar, H.H.V., 1978. Simulation of Assimilation, Respiration and Transpiration of Crops. Wageningen, The Netherlands.

Egli, D.B., Zhen-wen, Y., 1991. Crop growth rate and seeds per unit area in soybean. Crop Sci. 31, 439–442.

Evans, L.T., 1993. Crop evolution, Adaptation and Yield. Cambridge University Press, Cambridge.

Farquhar, G.D., von Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. Planta 149, 78–90.

Farrar, J.F., 1996. Sinks-integral parts of a whole plant. J. Exp. Bot. 47, 1273–1279.

Fehr, W.R., Caviness, C.E., 1977. Stages of Soybean Development. Iowa State University, Special Report 80, Ames, Iowa.

Forrester, J.W., 1961. Industrial Dynamics. MIT Press, Massachusetts Institute of Technology, and John Wiley and Sons, Inc., New York.

Goudriaan, J., 1986. A simple and fast numerical method for the computation of daily totals of crop photosynthesis. Agric. Forest Meteorol. 38, 249–254.

Harley, P.C., Weber, J.A., Gates, D.M., 1985. Interactive effects of light, leaf temperature, CO_2 and O_2 on photosynthesis in soybeans. Planta 165, 249–263.

Hunt, L.A., Boote, K.J., 1998. Data for model operation, calibration, and evaluation. In: Tsuji, G.Y., Thornton, P., Hoogenboom, G. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 9–39.

Irmak, A., Jones, J.W., Mavromatis, T., Welch, S.M., 2000. Evaluating methods for simulating soybean cultivar responses using cross validation. Agron. J. 92, 1140–1149.

Jagtap, S., Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. Agric. Ecosyst. Environ. 93, 73–85.
Janssen, P.H.M., Heuberger, P.S.C., 1995. Calibration of process-oriented models. Ecol. Model. 83, 55–66.
Jones, J.W., Hoogenboom, G., Wilkens, P.W., Porter, C.H., Tsuji, G.Y.E., 2003. Decision Support System for Agrotechnology Transfer Version 4.0: vol. 4. DSSAT v4: Crop Model Documentation. University of Hawaii, Honolulu, HI.
Kropff, M.J., van Laar, H.H., 1993. Modelling Crop–Weed Interactions. CAB International and International Rice Research Institute, Wallingford, Great Britain.
Mayaki, W.C., Teare, I.D., Stone, L.R., 1976. Top and root growth of irrigated and nonirrigated soybeans. Crop Sci. 16, 92–94.
Michalsky, J.J., 1988. The astronomical almanac's algorithm for approximate solar position (1950–2050). Solar Energy 40, 227–235.
Muchow, R.C., Sinclair, T.R., 1986. Water and nitrogen limitations in soybean grain production. II. Field and model analysis. Field Crop Res. 15, 143–156.
Pedersen, P., Boote, K.J., Jones, J.W., Lauer, J.G., 2004. Modifying the CROPGRO-Soybean model to improve predictions for the upper midwest. Agron. J. 96, 556–564.
Robinson, A.P., Conley, S.P., Volenec, J.J., Santini, J.B., 2009. Analysis of high yielding, early-planted soybean in Indiana. Agron. J. 101, 131–139.
Roder, W., Mason, S.C., Clegg, M.D., Kniep, K.R., 1989. Crop root distribution as influenced by grain sorghum-soybean rotation and fertilization. Soil Sci. Soc. Am. J. 52, 1337–1342.
Salvagiotti, F., Cassman, K.G., Specht, J.E., Walters, D.T., Weiss, A., Dobermann, A., 2008. Nitrogen uptake, fixation and response to fertilizer N in soybeans: a review. Field Crop Res. 108, 1–13.
Setiyono, T.D., Weiss, A., Specht, J., Bastidas, A.M., Cassman, K.G., Dobermann, A., 2007. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. Field Crop Res. 100, 257–271.
Setiyono, T.D., Weiss, A., Specht, J.E., Cassman, K.G., Dobermann, A., 2008. Leaf area index simulation in soybean grown under near-optimal conditions. Field Crop Res. 108, 82–92.
Sheldrake, A.R., 1979. A hydrodynamical model of pod-set in pigeon pea (*Cajanus cajan*). Indian J. Plant Physiol. 22, 137–143.
Sinclair, T.R., 1986. Water and nitrogen limitation in soybean grain production. I. Model development. Field Crop Res. 15, 125–141.
Sinclair, T.R., Horie, T., 1989. Leaf nitrogen, photosynthesis, and crop radiation use efficiency: a review. Crop Sci. 29, 90–98.
Sinclair, T.R., Seligman, N., 2000. Criteria for publishing papers on crop modeling. Field Crop Res. 68, 165–172.
Spaeth, S.C., Sinclair, T.R., Ohnuma, T., Konno, S., 1987. Temperature, radiation, and duration dependence of high soybean yields: measurement and simulation. Field Crop Res. 16, 297–307.
Specht, J., Hume, D.J., Kumudini, S.V., 1999. Soybean yield potential. A genetic and physiological perspective. Crop Sci. 39, 1560–1570.
Spitters, C.J.T., Toussaint, H.A.J.M., Goudriaan, J., 1986. Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis. Part I. Components of incoming radiation. Agric. Forest Meteorol. 38, 217–229.
Supit, I., Van der Groot, E., 2003. Updated System Description of the WOFOST Crop Growth Simulation Model as Implemented in the Crop Growth Monitoring System, CGMS, Applied by the European Commission. Treemail Publisher, Heelsum, The Netherlands.
Trybom, J., Jeschke, M., 2008. Foliar fungicide effect on soybean yield. Crop Insights 18, 1–4.
van Diepen, C.A., Wolf, J., van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. Soil Use Manage. 5, 16–24.
Verma, S.B., Dobermann, A., Cassman, K.G., Walters, D.T., Knops, J.M.H., Arkebauer, T.J., Suyker, A.E., Burba, G.G., Amos, B., Yang, H.S., Ginting, D., Hubbard, K.G., Gitelson, A.A., Walter-Shea, E.A., 2005. Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems. Agric. Forest Meteorol. 131, 77–96.
Vessey, J.K., Layzell, D.B., 1990. Regulation of assimilate partitioning in soybean. Plant Physiol. 83, 341–348.
Wang, E., Engel, T., 1998. Simulation of phenological development of wheat crops. Agric. Syst. 58, 1–24.
Wang, F., Fraisse, C.W., Kitchen, N.R., Sudduth, K.A., 2003. Site specific evaluation of the CROPGRO-soybean model on Missouri claypan soils. Agric. Syst. 76, 985–1005.
Watt, M., Evans, J.R., 2003. Phosphorus acquisition from soil by white lupin (*Lupinus albus* L.) and soybean (*Glycine max* L.), species with contrasting root development. Plant Soil 248, 271–283.
Williams, L.E., DeJong, T.M., Phillips, D.A., 1981. Carbon and nitrogen limitations on soybean seedling development. Plant Physiol. 68, 1206–1209.
Yin, X., Kropff, M.J., McLaren, G., Visperas, R.M., 1995. A nonlinear model for crop development as a function of temperature. Agric. Forest Meteorol. 77, 1–16.
Yin, X., van Laar, H.H., 2005. Crop System Dynamics. An Ecophysiological Simulation Model for Genotype-by-environment Interactions. Wageningen Academic Publishers, Wageningen, The Netherlands.