ELSEVIER

Physiology and modelling of traits in crop plants: implications for genetic improvement

K.J. Boote[a],*, M.J. Kropff[b], P.S. Bindraban[c]

[a]University of Florida, PO Box 110500, Gainesville, FL 32611-0500, USA
[b]Crop and Weed Ecology Group, Department of Plant Sciences, Wageningen University, PO Box 430,
6700 AK Wageningen, The Netherlands
[c]Plant Research International, Wageningen University and Research Centre, PO Box 16,
6700 AA Wageningen, The Netherlands

Abstract

Crop growth models have excellent potential for evaluating genetic improvement, for analyzing past genetic improvement from experimental data, and for proposing plant ideotypes for target environments. Crop models used for these plant breeding applications should be sufficiently mechanistic that processes can be investigated in a manner familiar to crop physiologists and plant breeders. In addition, the crop models must consider a sufficient number of cultivar-specific traits descriptive of life cycle phases, vegetative traits, and reproductive growth attributes. In this paper, we discuss how crop models consider genetic variability within a species (cultivar variation), how varietal characteristics can be determined from variety trial or other data, how crop models can be used to evaluate past genetic improvement, and how crop models can be used to hypothesize ideotypes for specific environments. We conclude that crop growth models can partially reproduce genotype by environment interactions when considered across broad ranges of weather and sites, and that crop models can be used to help plant breeders target cultivar improvement for specific environments. However, more physiological insight into primary processes such as source–sink relationships and morphological development will be needed for enhanced application of the models in breeding programmes. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Crop modelling; Plant breeding; Ideotype; Physiological traits; Crop improvement

1. Introduction

There is a general agreement among plant breeders and physiologists that physiological knowledge can be applied to improve breeding efficiency (Jackson et al.,

---

* Corresponding author.
   E-mail address: kjb@gnv.ifas.ufl.edu (K.J. Boote).

1996); however, the contribution has so far been small to moderate. Using empirical selection, breeders have been successful in increasing yield potential in the past; however, progress in yield potential is increasingly difficult to maintain and in some crops such as rice no progress in increasing yield potential has been achieved in the past decades. There is interest in taking an ideotype approach to design plants for target environments, to take advantage of interactions of plant characteristics and the environment. Shorter et al. (1991) proposed collaborative efforts between breeders, physiologists and modellers, using simple biological models as a framework to integrate physiology with breeding and to evaluate adaptation of genotypes to target environments. While current crop models still lack the ability to describe all the subtle complexities associated with genotypic differences, they have been powerful in unravelling mechanisms that determine crop yield in relation to the environment (Bindraban, 1997). In addition, Hammer et al. (1999) proposed that crop models based on physiologically sound mechanisms, can quantify and integrate crop responses to genetic, environmental, and management factors, and thus, have the potential to predict performance of individual genotypes in different managements and environments.

The first objective of this paper is to discuss how crop growth simulation models have addressed cultivar variation in physiological/genetic traits and how such cultivar traits have been derived in the past. The second objective is to illustrate how the models can be used to evaluate genetic improvement and hypothesize ideotypes for specific environments. The focus is on trait physiology and using physiological insight to explain genetic differences that contribute to yield.

2. Modelling cultivar differences

Plant breeders would agree that genetics is the basis for differences between species as well as differences between cultivars within the same species. The distinction is primarily a matter of degree of genetic variability. But how well do crop models presently account for genotypic (cultivar) variation? We believe there is an inadequate degree of genotypic specificity in present crop growth models for several reasons: (1) models are relatively new; (2) crop modellers are mainly working on the more basic problems of accurately modelling the crop's response to climate, soil, management, and their interactions; (3) many models are relatively generic (and too simplistic); and (4) most lack the ability to describe the subtle complexities associated with genotypic differences within species. Nevertheless, many crop modellers would agree with the need and potential ability of crop models to address cultivar differences within a species.

2.1. Species differences

Crop models have been developed for a number of different species, and the characteristics that make each species unique, are often buried in the programming code. Thus, it is difficult to distinguish between real species differences and the

particular approach used by different programmers. In some cases, such as with the CROPGRO model and the SUCROS-type models, the same FORTRAN code is used for a number of different crops, and species characteristics are contained in a data file. For CROPGRO, the 'species' file describes basic relationships of photosynthesis, growth and maintenance respiration, N_2 fixation, carbohydrate and N mobilization, partitioning, vegetative growth (leaf area, height, width), reproductive growth (addition and growth of seeds, pods), etc. to climatic and soil conditions as well as plant internal water and N status. Readers are referred to Boote et al. (1998), Kropff and Van Laar (1993), and Goudriaan and Van Laar (1994) for more information on how species coefficients for the physiological processes are used in the CROPGRO and SUCROS-type models. For example, the species file describes leaf photosynthesis response to temperature, carbon dioxide, photon flux density, and leaf N. It describes the temperature and plant water status effects on nodule growth rate and specific nitrogenase activity. It describes the basic temperature effects on rates of pod and seed addition, pod growth, and seed growth. Keeping these process sensitivities in a species file intended for use only by model developers, allows a smaller set of physiological (genetic) traits for cultivar/ecotype differences that are easier for model users to understand and describe. This shorter list of cultivar traits focuses on heat units or physiological days for different life cycle phases, critical daylength parameters, key vegetative traits such as determinacy of leaf area growth, determinacy of pod addition, seed size, seed fill duration, seed oil and protein, etc.

2.2. Cultivar coefficients (within a species)

The various crop models have different numbers of cultivar coefficients that describe how a cultivar divides up its life cycle into phase durations. There are phase modifiers (critical daylength, daylength sensitivities, vernalization requirements), vegetative attributes, and reproductive attributes. Here, we list and define cultivar coefficients presently used for the CROPGRO model (Table 1). For description of cultivar coefficients used with the CERES models, see Hoogenboom et al. (1994), Appendix A, p. 201 and p. 208. See Yin et al. (1997) for description of the coefficients used in ORYZA1 (SUCROS-Rice model).

2.3. Variation contributed by cultivar traits in the CROPGRO model

In order to better illustrate the practical meaning of cultivar coefficients defined in Table 1 for CROPGRO, we show a comparison of different maturity groups within soybean and discuss their simulated life cycles and yield in Florida and Iowa (Table 2). Soybean cultivars within the USA have been broadly classified into maturity groups (MG) from 000 to 12, based primarily on their sensitivity to daylength which influences their life cycle duration.

Soybean cultivars in MG 00 have shorter basic life cycles and are less daylength sensitive. These are grown in northerly latitudes in the USA, Canada, and Europe. By contrast, cultivars grown in the southern USA and in the tropics have longer basic life cycles and greater daylength sensitivity, approaching MG 12 near the

Table 1
Cultivar coefficients for the CROPGRO model, value ranges typical for soybean

Trait	Range	Definition of trait
ECO#	SB0301	Code for the ecotype to which this cultivar belongs (see *.eco file)
CSDL	14.6–11.8	Critical short day length below which reproductive development progresses with no daylength effect (for short day plants) (h)
PPSEN	0.129–0.349	Slope of the relative response of development to photoperiod with time (positive for short day plants) (h^{-1})
EM–FL	15.5–23.5	Time between plant emergence and flower appearance (R1) (photothermal days)
FL–SH	5.0–10.0	Time between first flower and first pod (R3) (photothermal days)
FL–SD	12.0–16.0	Time between first flower and first seed (R5) (photothermal days)
SD–PM	27.0–38.0	Time between first seed (R5) and physiological maturity (R7) (photothermal days)
FL–LF	15.0–26.0	Time between first flower (R1) and end of leaf expansion (photothermal days)
LFMAX	0.92–1.17	Maximum leaf photosynthesis rate at 30 C, 350 vpm CO2, and high light (mg CO2 m^{-2} s^{-1})
SLAVR	355–400	Specific leaf area of cultivar under standard growth conditions (cm^2 g^{-1})
SIZELF	140–200	Maximum size of full leaf (three leaflets) (cm^2)
XFRT	1.00	Maximum fraction of daily growth that is partitioned to seed + shell
WTPSD	0.15–0.22	Maximum weight per seed (g)
SFDUR	20–25	Seed filling duration for pod cohort at standard growth conditions (photothermal days)
SDPDV	1.9–2.5	Average seed per pod under standard growing conditions (No. pod^{-1})
PODUR	8.0–14.0	Time required for cultivar to reach final pod load under optimal conditions (photothermal days)
Frequently used important traits from ECO file		
R1PRO	0.189–0.549	Increase in daylength sensitivity after anthesis (CSDL decreases by this amount) (h)
FL–VS	6.0–27.0	Time from first flower to last leaf on main stem (photothermal days)
THRESH	76–79	The maximum ratio of [seed/(seed + shell)] at maturity. Causes seed to stop growing as their dry weight increases until shells are filled in a cohort.
SDPRO	0.37–0.43	Fraction protein in seeds [g (protein) g^{-1} (seed)]
SDLIP	0.18–0.22	Fraction oil in seeds [g (oil) g^{-1} (seed)]

equator. Table 2 shows critical short daylength values (CSDL), photoperiod sensitivity slopes (PP-SEN), and physiological day (PD) requirements to complete life cycles for MG 00-9 soybean cultivars as solved by Grimm et al. (1993, 1994) from experimental data.

Simulated days to maturity and yield performance are shown in Table 2 for MG 00–9 soybean cultivars grown in Florida and Iowa. The low-number MGs mature much too early in Florida and yield poorly. The optimum MGs in Florida for yield and most effective use of the season are MGs 6–9 (both in simulations and in actual production practice). In Iowa, the low-number MGs have a longer simulated life cycle and higher yield potential because of longer daylength and cooler temperature at that location. The optimum MGs for Iowa for yield and effective use of the season are MGs 2–3, again corresponding to actual practice. MG 4 was reasonably

Table 2
Effect of soybean maturity group (MG) and genetic coefficients on days to maturity and yield for crops grown under rainfed conditions in Gainesville, Florida and Ames, Iowa[a]

MG	Cultivar parameters used in model					Florida simulation		Iowa simulation	
	CSDL (h)	PPSEN (day h^{-1})	EM–FL (PD)	FL–SD (PD)	SD–PM (PD)	Maturity (days)	Yield (kg ha^{-1})	Maturity (days)	Yield (kg ha^{-1})
00	14.35	0.148	16.0	12.0	30.0	70.0	834	93.1	1936
0	14.10	0.171	16.8	13.0	31.0	73.3	900	100.6	2295
1	13.84	0.203	17.0	13.0	32.0	76.7	971	108.7	2625
2	13.59	0.249	17.4	13.5	33.0	83.8	1119	121.1	3072
3	13.40	0.285	19.0	14.0	34.0	93.6	1411	134.8	3220
4	13.09	0.294	19.4	15.0	34.5	107.3	1971	151.5[b]	3087
5	12.83	0.303	19.8	15.5	35.0	119.5	2326	164.0[c]	2375
6	12.58	0.311	20.2	16.0	35.5	132.2	2678	–[d]	
7	12.33	0.320	20.8	16.0	36.0	145.0	2741	–[d]	
8	12.07	0.330	21.5	16.0	36.0	159.0	2846	–[d]	
9	11.88	0.340	23.0	16.0	36.5	170.0	2716	–[d]	

[a] Crops were planted on day 123 at 30 plants m^{-2} and grown under rainfed conditions, using 10 years of historical weather at each site (1978–1987 in Florida, and 10 of 1984–1995 in Iowa). CSDL, critical short daylength; PPSEN, photoperiod sensitivity; PD, physiological days; EM–FL, emergence to flowering; FL–SD, flowering to first seed, and SD–PM, first seed to physiological maturity, are given for each MG, as used by the CROPGRO-soybean model. Physiological day requirement to emergence is 3.6 days.

[b] Freeze damage in 2 of 10 years.

[c] Freeze damage in 5 of 10 years.

[d] Freeze damage in 10 of 10 years and maturity not reached.

high in yield, but suffered freeze damage in two of 10 seasons and MG 5 suffered freeze damage in 5 of 10 years. MGs 5–12 are not suited for Iowa because they have shorter critical daylength and greater daylength sensitivity, which prevents them from completing their life cycle prior to frost. While the CSDL and PP-SEN are the dominant traits affecting cultivar fit to latitude zones, notice that the basic physiological day requirement for emergence to flowering (EM–FL), flowering to beginning seed (FL–SD), and beginning seed to physiological maturity (SD–PM), all increase with successively higher MG classes.

While Table 2 illustrates a reduced set of five traits (of the 15) from the CROP-GRO cultivar file, some of the other traits are somewhat proportional to these five. For example, the single seed-filling duration (SFDUR), at least for soybean, is approximately three quarters of the value used for SD–PM. Likewise, the time to the end of leaf area expansion and end of main stem node expression is dependent on the predicted time to flowering, plus an additional phase (FL–LF or FL–VS) that differs with determinant versus indeterminate cultivar type.

2.4. Variation contributed by cultivar traits in CERES model

CERES-Maize has only five coefficients for cultivar traits. Table 3 compares the life cycle and simulated performance of nine maize hybrids typical of those in the

Table 3
Effect of maize variety selection and modification of varietal coefficients on days to anthesis, days to maturity and grain yield for crops grown rainfed under 10 years of weather at Gainesville, Florida (1978–1987) and Ames, Iowa (1980–1989)[a]

Variety (hybird)	P1	P2	P5	G2	G3	G'ville, FL, 1978–1987			Ames, Iowa 1980–1989		
						Anthesis (days)	Maturity (days)	Yield (kg ha^{-1})	Anthesis (days)	Maturity (days)	Yield (kg ha^{-1})
Standard hybrids											
CP170	120	0.00	685	908	10	60.5	99.8	3143	55.8	100.3	8282
F478×W705A	140	0.00	685	908	10	63.1	101.6	3305	59.2	103.5	8672
F16×F19	165	0.00	685	908	10	65.9	106.5	3618	62.3	107.1	8817
DeKalb XL45	150	0.40	685	908	10	65.6	103.6	3624	63.7	108.8	9036
B14×C103	180	0.50	685	908	10	69.2	109.0	3946	67.7	113.2	8850
A632×VA26	240	0.30	685	908	10	77.8	116.0	4860	73.7	120.6	8734
PIO 3147	255	0.76	685	908	10	79.3	117.5	4833	77.0	124.7	8812
H6	310	0.30	685	908	10	85.7	122.9	4569	82.5	132.1	8992
H. OBREGON	360	0.80	685	908	10	91.4	128.1	4964	92.9	147.6	9717
Sensitivity test											
P5 = 685	240	0.30	685	908	10	77.8	116.0	4860	73.7	120.6	8734
P5 = 785	240	0.30	785	908	10	77.8	121.5	5706	73.7	128.8	10181
P5 = 885	240	0.30	885	908	10	77.8	126.8	6615	73.7	138.4	12112
G2 = 908	240	0.30	685	908	10	77.8	116.0	4860	73.7	120.6	8734
G2 = 808	240	0.30	685	808	10	77.8	116.0	4549	73.7	120.6	8209
G2 = 708	240	0.30	685	708	10	77.8	116.0	4183	73.7	120.6	7377
G3 = 10	240	0.30	685	908	10	77.8	116.0	4860	73.7	120.6	8734
G3 = 8	240	0.30	685	908	8	77.8	116.0	4117	73.7	120.6	7360
G3 = 6	240	0.30	685	908	6	77.8	116.0	3124	73.7	120.6	5724

[a] Crops were sown at 7.2 plants m^{-2} in 0.91-m rows on 2 February (Florida) and 2 May (Iowa) and grown under rainfed conditions on a Millhopper fine sand (Florida) or a Nicollet silty clay loam (Iowa).

CERES-maize cultivar file, when grown at Gainesville, FL or Ames, IA. P1 defines the basic vegetative phase prior to photoperiod sensitivity (degree-days, using Tb of 8°C). P2 is the daylength sensitivity coefficient [the extent (days) that development is delayed for each hour increase in photoperiod above the longest photoperiod (12.5 h) at which development proceeds at maximum rate]. Together P1 and P2 determine the length of time to anthesis. With all other traits unchanged, life cycle (anthesis and maturity date) increases with increasing P1 and P2 and yield potential increases up to 116-day life cycle in Florida and 108-day life cycle in Iowa. This is related to optimizing lower leaf area index (LAI) and vegetative mass prior to grain fill. The much lower grain yield in Florida than in Iowa is attributed to greater water deficit from low springtime rainfall and sandy soil. P5 is the grain filling duration (degree-days from silking to physiological maturity). Sensitivity analysis shows that increase in P5 increases yield considerably at both sites (12–13% yield increase per 10% increase in P5). G2 is the maximum number of kernels per plant, while G3 is the kernel filling rate in mg kernel^{-1} day^{-1} under optimum conditions. Decreasing G2 from its typical value of 908 kernels per plant decreases yield 6–7% per 10% decrease

in G2 in Florida and Iowa. Sensitivity to changes in kernel filling rate (G3) are quite dramatic and linear, with 8–9% yield decrease per 10% change in G3. The use of G2 and G3 allows CERES-maize users to set grain yield for their particular hybrids at their sites, but this strong linear responsiveness to kernel numbers and kernel growth rate causes problems because the numbers of kernels per plant is quite environmentally dependent and does not follow so predictably in other seasons.

By comparison, the CERES-wheat model uses relative scaling (from 0.0 to 6.0) for wheat cultivar coefficients, except for the kernel number per unit stem weight (G1) and kernel filling rate (G2). The CERES-wheat model has a vernalization coefficient (6.0 for winter wheats versus 0.5 for spring wheats) and has a photoperiod coefficient that slows development if daylength is less than 20 h. Similar to CERES-maize, the radiation use efficiency coefficient for CERES-wheat is hard-coded and constant for all cultivars.

2.5. Variation contributed by cultivar traits in ORYZA1 rice model

For rice, Yin et al. (1997) determined the optimal preflowering phenology in irrigated rice for several contrasting environments using the model ORYZA1 (Kropff et al., 1994). The model was extended with a detailed phenology model, and yield optimum was determined as a function of duration of the vegetative phase. Five coefficients were distinguished: the basic preflowering period, the optimum day temperature, the optimum night temperature, the temperature sensitivity and the photoperiod sensitivity (Yin et al., 1997). Environments ranged from tropical wet seson/dry season to subtropical and temperate. For each environment there was an optimal preflowering period that produced the highest yield (Fig. 1). In the subtropics, rice yield potential was restricted by the available growing season (temperature window). In the tropics, a critical preflowering period was found beyond which yield potential did not increase. The simulated optimal pre-flowering periods in these environments matched the pre-flowering period of the best varieties bred in those environments. This indicates that models can be used to target phenological traits of genotypes for specific environments.

From this discussion, it is obvious that modellers have different ways of modelling cultivar differences. While one may argue whether these have really captured the genetic differences, it is apparent that the approach and degree of detail more likely match what the modellers saw as differences in real data, and represent mathematical coefficients resulting from solving from data sets. Some modellers have proposed the notion of emergent properties for genetic coefficients in crop models (Hammer, 1998; De Wit and Penning deVries, 1983). This is a worthy objective, if well-grounded and tested in outcomes.

3. Solving for cultivar traits — intensive versus extensive data

What can we do if our goal is to adequately describe new cultivars, so we can use the models to recommend site-specific management practices for producers? Must
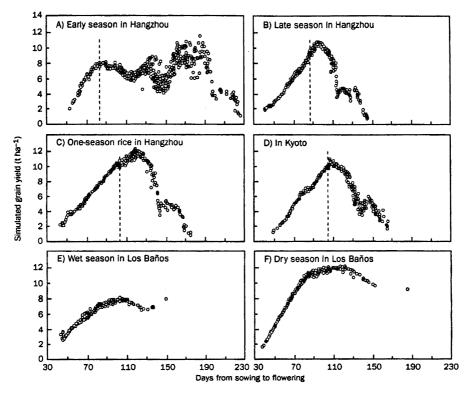
Fig. 1. Relation between simulated grain yield and preflowering period in rice for different growing seasons in three locations. Vertical dashed lines indicate the maximum preflowering period as determined by the length of the available growing season (after Yin et al., 1997).

we obtain intensive crop analyses data, or can we develop alternative procedures that are less demanding of data, for example, using only variety trial information? For the intensive approach, crop modellers have used fairly detailed information on final yield attributes, crop vegetative and reproductive growth stages during the season, plus periodic in-season sampling for dry matter accumulation and partitioning among plant components. (See later section where such detailed growth data was used to solve genetic coefficients from old and new cultivars to evaluate genetic yield improvement.) Such detailed information is reasonable for testing of relatively new crop models, especially for new locations. However, alternate approaches are needed, to access information available from yield trials conducted on large numbers of cultivars. It is not feasible or cost-efficient to ask for detailed growth analyzes on the hundreds of new varieties that become available each year. Even worse, the information would always be several years out-of-date, even if resources for massive growth measurements were provided.

We have developed a procedure for solving cultivar traits for the CROPGRO-soybean model from extensive variety trials conducted by states, private companies, and regional testing mechanisms (Mavromatis et al., 2001). Typical data collected

include: harvest maturity, seed yield, seed size, height, lodging, and sometimes flowering date. Weather data, soil, and management information are also required, because the procedure uses the CROPGRO model in an optimization shell in conjunction with the experimental data. We tested this procedure with variety trial data for a set of 11 cultivars grown at five sites in Georgia over 4–10 years (Irmak et al., 2000; Mavromatis et al., 2001). It is best to have data available from a minimum of 20–30 site-year combinations, such as 3 years over seven sites differing in latitude and planting date.

In our sequential procedure, cultivars were first assigned to a default MG as closely as known. Then, for each cultivar across all sites and years, we solved for CSDL and PP-SEN needed to predict observed anthesis date. Then, we solved for the duration from anthesis to maturity (FL–SD + SD–PM) as well as a coefficient, R1PRO, that decreases CSDL by an offset, thus making the crop more sensitive to daylength after anthesis. Piper et al. (1996) provided evidence that soybean cultivars had increased sensitivity to daylength after anthesis (decreasing CSDL), particularly for higher number MGs. Fig. 2 compares predicted versus observed maturity of the Stonewall cultivar after these two optimization steps. The RMSE of prediction of anthesis and maturity was 3.1 and 6.4 days, respectively.

Secondly, for each site, we solved for site indices (fertility and water holding capacity) across all cultivars and years, with site-year average yield as the objective function. The goal here was to remove variation attributed to site-specific factors of fertility and water that were common across cultivars, prior to solving for remaining yield differences attributed to cultivar variation in yield within site.

Next, for each cultivar we used a two-way linear grid search that optimized on yield differences between cultivars based on two types of traits such as: (1) light-saturated rate of leaf photosynthesis LFMAX; to increase/decrease yield, and (2) a shift in fraction of time from anthesis to maturity devoted to seed fill (i.e. increase
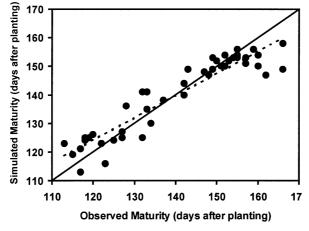

Fig. 2. Simulated versus observed maturity, after solving traits of Stonewall soybean cultivar from Georgia variety trial data (Mavromatis et al., 2001). The 1:1 line and regression line are shown.

yield by shorter FL–SH and FL–SD and shorter duration of rapid pod addition (PODUR), but longer SD–PM and longer seed filling duration (SFDUR)). These two traits have different pathways for affecting yield, because the first acts primarily on amount of biomass and the latter affects seed harvest index (HI). We allow either of these traits to account for yield variation among cultivars; however, if cultivar trials were able to provide data on final seed HI, we could select for one trait over the other. With the availability of growth analyses data, one could distinguish between photosynthesis-enhancing versus HI-shifting traits. Knowledge of pedigree and genetics among the cultivars could help focus particular traits.

With this optimization procedure, the RMSE for yield of Stonewall, a MG 7 cultivar, was decreased from 769 kg ha^{-1} starting with the default general MG, to 602 kg h^{-1} after solving phenology traits, to 414 kg ha^{-1} after solving for phenology and site indices, to 404 kg ha^{-1} after optimizing for phenology, site, and cultivar yield potential traits. Fig. 3 shows a 1:1 plot of predicted versus observed yield for the Stonewall cultivar. Site effects accounted for more of the total variation in yield than did cultivar effects even when phenology (life cycle) was included among cultivar traits, but perhaps this is not surprising to field agronomists because Georgia had considerable yield variation in these years caused by drought stress. This procedure also succeeded in predicting some of the G×E variability observed in these trials.

The DSSAT software contains a helpful tool, GENCALC, that has been used to solve for physiological coefficients from field data sets (Hunt et al., 1993). The principle of GENCALC is basically a linear search on successive coefficients against single values in the field performance file, and was designed for one experiment and one variety at a time. Coefficients are then averaged over multiple experiments. The optimization software described earlier by Mavromatis et al. (2001) has the advantage of handling larger data sets, with sophisticated search techniques on more than one coefficient simultaneously, based on linear grid search or simulated annealing.
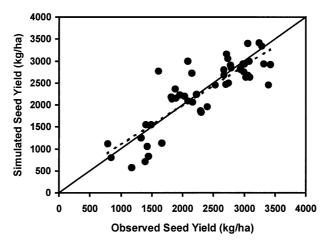

Fig. 3. Simulated versus observed seed yield, after solving traits of Stonewall soybean cultivar from Georgia variety trial data (Mavromatis et al., 2001). The 1:1 line and regression line are shown.

Care must be taken in the above optimization methods, to avoid blindly following the optimization. In such optimization, it is important to consider knowledge of what the crop and the model are doing. Knowledge of physiology and how the crop model uses that physiology must be considered carefully, and in the correct sequence, and with the appropriate weighting relative to feasible genetic range for individual traits.

4. Using crop models to evaluate mechanisms of past genetic yield improvement

Crop models can be used in a dynamic framework to generate hypotheses and to investigate causal relationships, such as reasons for weather and management effects on yield, or genetic causes for yield improvement. Here, we give an example of how the CROPGRO-soybean model was used in an inverse engineering mode to evaluate hypothetical physiological mechanisms associated with past genetic yield improvement. Such evaluations require detailed data on growth and processes. Genetic coefficients can also be derived from growth analyses data in this manner. In a study funded by the Iowa–Illinois soybean promotion boards, detailed growth analyses data were collected on two new cultivars (Kruger 2828 and Stine 3660) and one old cultivar (Williams 82) at Lewis, IA (Batchelor, Kishore, Whigham, and Boote, unpublished). The trials were conducted in environments not limited by biotic pests. In addition, soil water measurements were collected to be sure the model's soil water balance was working correctly against observed soil water measurements. The soil fertility parameter (SLPF) was adjusted to predict the dry matter gain averaged over all cultivars being compared, to avoid attributing site differences to cultivar traits. Then we solved for cultivar traits, initially setting phenology/life cycle traits, followed by traits affecting partitioning, determinacy, photosynthesis, and N mobilization. Cultivar coefficients of daylength sensitivity (CSDL and PP-SEN) and phase durations (EM–FL, FL–SH, FL–SD, and SD–PM) were used to set cultivar differences in life cycle and the onset of pod and seed additions, as well as the phenological time from beginning seed to physiological maturity. Cultivar differences in determinacy of pod set were apparent, and were set by varying the length of PODUR. Cultivar differences in dry matter accumulation during mid-life cycle were adjusted by increasing or decreasing LFMAX.

Fig. 4 compares crop and seed growth of new cultivars Kruger 2828 and Stine 3660 to the older, lower-yielding Williams 82 cultivar, after the above calibration process. Table 4 shows the simulated responses over 2 years for two sites. Simulated K-2828 and S-3660 had 15–19% higher yields, earlier podset (1–7 days), earlier seedset (3–8 days), longer time from seed to maturity (4–5 days), and higher harvest index than Williams 82. Model traits that had to be changed for the new cultivars included: 30% faster pod addition, 8–11% longer SD–PM, 7–9% longer SFDUR, 9% higher leaf photosynthesis, and 10% slower leaf N mobilization. The earlier and more rapid seed growth (Fig. 4) and the higher seed harvest indices (Fig. 5, Table 4) are modelled outcomes of changing these traits for the new cultivars. While higher photosynthesis contributed to greater yields, much of the yield gain also came from
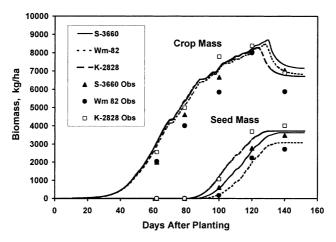
Fig. 4. Comparison of simulated and observed growth dynamics for seed and total crop mass of new soybean cultivars Kruger 2828 and Stine 3660 compared to old cultivar Williams 82 at Lewis, Iowa in 1997. From Batchelor, Kishore, Whigham, and Boote (unpublished, 1999).

improved partitioning allowed by greater determinacy of seedset and longer time allocated to seed-fill as shown by the improved seed harvest index (Fig. 5). Final simulated harvest index was increased from 0.493 for Williams-82 to 0.539 for Stine 3660, and 0.589 for Kruger 2828 (Table 4). The new cultivars had higher dry matter accumulation during mid-life cycle than Williams 82, which required a LFMAX value of 1.08 mg CO_2 m^{-2} s^{-1} for the two new cultivars compared with 0.99 mg CO_2 m^{-2} s^{-1} for Williams 82. In addition, the new cultivars had more sustained dry matter accumulation during seedfill, which we were able to mimic by 10% slower rate of N mobilization. Results from three sites in Illinois for comparisons of other sets of new versus old cultivars in the same project strongly confirm similar changes in cultivar traits. These findings nicely illustrate how crop models can be used to analyse past genetic improvement, and point the way for further genetic improvement. This also illustrates the technique of solving for genetic traits by changing modelled traits by comparison to in-season growth analyses.

5. Applications of models to hypothesize yield improvement

Models have excellent potential for hypothesizing yield-improvement possibilities based on improving single or multiple combinations of genetic traits. Model sensitivity analyses of genetic traits have been conducted for a variety of crops by Duncan et al. (1978), Landivar et al. (1983a, 1983b), Wilkerson et al. (1983), Boote and Jones (1986), Whisler et al. (1986), Elwell et al. (1987), Boote and Tollenaar (1994), Kropff et al. (1995), Hammer et al. (1996), Aggarwal et al. (1997), and Boote and Jones (1998). The value of such analyses depends in part upon the realism with which the models mimic the various physiological processes and traits, and the

Table 4
Simulated phenology dates, seed yield, and seed harvest index for comparison of new cultivars, K-2828 and S-3660, versus old cultivar Williams 82 at two sites in Iowa in 1997 and 1998 (growth analyses data courtesy of Batchelor, Kishore, Whigham, and Boote)

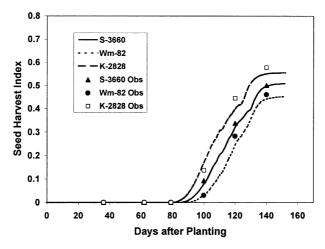| Cultivar	MG	R1 (days)	R3 (days)	R5 (days)	R7 (days)	Yield (kg ha^{-1})	Seed HI (fraction)
Kruger-2828	2.7	59	73	82	126	4264	0.589
Stine-3660	3.6	65	79	87	131	4125	0.539
Williams-82	3.8	66	80	90	130	3590	0.493

Fig. 5. Comparison of simulated and observed seed harvest index (HI) of new soybean cultivars Kruger 2828 and Stine 3660 compared with old cultivar Williams 82 at Lewis, Iowa in 1997. From Batchelor, Kishore, Whigham, and Boote (unpublished, 1999).

extent to which the models contain natural feedbacks (honoring C, N, and water balance) and mimic known pleiotropic responses (increased leaf photosynthesis coming partly from increased leaf thickness, but lesser leaf area). In addition, for a realistic response it is important to vary traits only within the range of known genetic variation. Boote and Tollenaar (1994) reported that simulated response to variation in leaf photosynthesis is strongly dependent on the degree of linkage of light-saturated leaf photosynthesis to specific leaf weight (SLW). With a strong linkage of maximum leaf photosynthesis (Pmax) to (SLW) taken from Dornhoff and Shibles, 1970), there is a clear optimum SLW (and leaf photosynthesis rate) at which yield was optimum. Overall, these simulations showed rather modest yield responses to increased light-saturated leaf rate, thus downplaying the great potential that physiologists proposed for leaf photosynthesis in the decades of the 1960s, 1970s, and 1980s.

Crop simulation can provide a quantitative evaluation of the value of a given trait in a given target environment only if the physiological basis of the trait action is

known and if physiological connections and required feedbacks are present (Hammer et al., 1996, 1999). These limitations should caution us that research into physiological mechanisms should continue until we are quite confident about how these operate. Bindraban (1999) suggested that correct representative description of physiological processes in crop models is critically important. For example, the hypothetical linear increase of photosynthesis with leaf nitrogen content would suggest a substantial increase in yield with increased leaf nitrogen content. An observed saturation response of photosynthesis to leaf nitrogen content, however, drastically reduces this scope of yield increase (Bindraban, 1999).

It is obvious that models differ considerably in the amount of complexity allowed for cultivar traits, as we have seen in Tables 1–3 and with example simulations using typical cultivar traits. Nevertheless, we can only vary the traits that are presently used in a given model. We must be cautious in our sensitivity analyses and interpretations because the models may fail to consider known cultivar variations, feedbacks, and pleiotropic effects. In CROPGRO, we have the possibility to illustrate life cycle shift (vary CSDL), to increase seed filling period and harvest index (SD–PM and SFDUR), to vary determinacy (FL–VS, FL–LF, and PODUR), to vary photosynthesis (LFMAX), specific leaf area, leaf N concentration, rate of N mobilization, seed protein, seed oil, rooting profile, early vegetative vigour, and cool temperature tolerance (not a presently cultivar trait, but Tb shift in species).

5.1. Increasing filling period to increase yield and seed HI

It may not be desirable to dramatically change total life cycle duration, because life cycle is what fits a cultivar to its particular latitude zone or target environment, particularly where frost or rainy season dictates the possible length of season. Cultivars adapted to a region are likely already nearly optimized for life cycle duration. Within the total life cycle, small changes in daylength sensitivity and shifts in allocation of life cycle to different phases are possible.

If we hold total life cycle unchanged, we can hypothetically increase the fraction of life cycle allocated to reproductive growth and decrease the proportion of time for vegetative growth. In CROPGRO this is done by decreasing time to anthesis (EM–FL), beginning pod (FL–SH), and beginning seed (FL–SD), while increasing SD–PM and SFDUR. We tested these ideas with CROPGRO-soybean, starting with standard coefficients for the Williams 82 cultivar at Ames, Iowa, with 11 years of relatively non-stressed weather data between 1980–1996. Soybean yield initially increases rapidly with increasing pod fill duration, but the response slows and becomes asymptotic at long filling periods (Fig. 6). The saturating or limiting effect on yield is associated with: (1) lower leaf area index (LAI) caused by shorter vegetative phase and (2) progressively lower leaf N concentration during seed fill which decreases photosynthesis and results in less N available to mobilize during seed-fill. Simulated maximum LAI declined steadily from 7.2 to 3.8 with increasing fraction of life cycle allocated to reproductive growth over the range of change shown in Fig. 6, thus decreasing light interception and photosynthesis. Likewise, simulated main stem node number and canopy height decreased 25 and 23%, respectively.

Creating 20% slower N mobilization increased the yield response considerably at longer filling periods, but not at short filling periods (Fig. 6).

Breeders, in selecting for higher yield, probably have already selected for slow N mobilization (more stay-green) along with the longer filling period. The simulated yield response per day increase from R5 to R7 (beginning seed to physiological maturity) is 27.4 and 36.8 kg ha^{-1} day^{-1} for normal N mobilization, and 20% slower N mobilization, respectively. This agrees closely to reported yield increase of 35 kg ha^{-1} for each day increase in time from R4 to R7, averaged over 119 cultivars at 10 site-years in the midwestern USA (Dunphy et al., 1979). There appears to be a genotype by environment interaction with the longer filling period trait. In the six more drought-prone years between 1980–1996 at the same site, longer filling period had no advantage over the mid-range standard, and was less advantageous because earlier podset resulted in less simulated root growth, less soil water extraction, more water stress, and greater drought-induced reductions of biomass and seed yield.

The above model simulations also show how the long-honored 'outcome' trait of, is actually not a trait at all, but is the natural outcome of factors that shift the timing of allocation of assimilate to seed growth (Fig. 7). Significantly, seed HI increases asymptotically with increased reproductive duration, and appears to achieve an asymptote at about 0.6. Note, however, that other traits, such as slower N mobilization, appear to push the asymptote up somewhat, whereas water deficit limits HI. Within the reasonable range about the standard Williams 82, an 11.3-day increase in reproductive duration increased HI from 0.434 to 0.490.

In the above sensitivity, there was no consideration for the genetic feasibility to obtain long seed filling periods. In order to interpret and evaluate the value/effect of any given cultivar trait, it is important to know the feasible genetic range for the
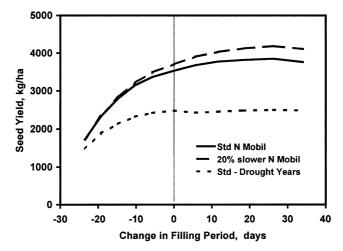

Fig. 6. Simulated soybean seed yield response to variation in the filling period duration (beginning seed, R5, to physiological maturity) at Ames, Iowa for rainfed weather (1980–1996). Treatments include standard N mobilization and 20% slower N mobilization (both with 11 adequate rainfall years) and standard N mobilization (with 6 drought-prone years).
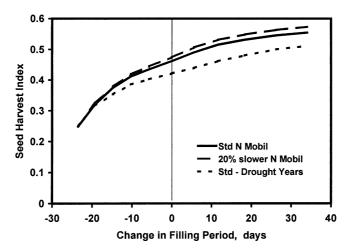
Fig. 7. Simulated soybean seed harvest index in response to variation in the filling period duration (beginning seed, R5, to physiological maturity) at Ames, Iowa for rainfed weather (1980–1996). Treatments include standard N mobilization and 20% slower N mobilization (both with 11 adequate rainfall years) and standard N mobilization (with 6 drought-prone years).

given trait. Boote and Tollenaar (1994) in an earlier SOYGRO simulation of pod filling period, summarized literature of Egli et al. (1984), Swank et al. (1987), and others, and concluded that it was genetically feasible to vary filling period by −10 and +10 days about the mean used for Williams 82. A simulated increase of 11.3 days centered over the Williams 82 standard trait, increased yield by 243 kg ha^{-1} (6.9%). Kropff et al. (1995) and Aggarwal et al. (1997) also indicated that a longer grain filling period is an essential trait for increasing the yield potential in rice. However, in preliminary experiments they found that the genetic variability in the trait is not large and there appeared to be strong negative trade-offs between grain filling duration and seed filling rate.

5.2. Determinate versus indeterminate: duration of leaf area expansion, node expression and rate of pod addition

Crop species and cultivars vary in degree of morphological and physiological determinacy (how rapidly and completely they make the transition from vegetative to reproductive growth). The end of leaf area expansion and end of main stem node expression is controlled substantially by the time to flowering, plus an additional phase that differs with determinant versus indeterminate cultivar type. In addition, determinacy frequently includes how rapidly pods are added, which is modelled by the PODUR trait, the physiological time required from start to end of rapid pod addition. We have had experience with coefficients for determinant cultivars in the South, indeterminate cultivars in the Midwest, as well as cultivars in MG 4 and five in the Mississippi delta, where both determinate and indeterminate cultivars are

common. The physiological days from first flower to end of main stem node appearance (FL–VS) is much less for determinate cultivars than for indeterminate cultivars (6–9 versus 26), thus resulting in fewer main stem nodes, and shorter plants. Internode length for determinates is about 90% of normal. Physiological days from flowering to end of leaf area growth (also on branches), FL–LF, is about 18 versus 26 for determinant versus indeterminate, causing somewhat lower LAI. For the example simulations in Table 5, we assumed the same dates of flowering, podset, and maturity, but allowed that determinant cultivars have 33% faster pod addition (8 versus 12 physiological days), shorter FL–LF (18 PD), and shorter FL–VS (9 PD). Simulating the above three changes together as determinacy (Table 5), we obtained 14.9 versus 20.3 main stem nodes, 79 versus 108 cm height, and 9.4% decrease in LAI in Iowa. Seed yield increased 2.4% and HI increased from 0.447 to 0.466, primarily because pods were added more rapidly.

5.3. Increasing leaf photosynthesis to increase yield

Leaf photosynthesis is a fairly conservative trait among adapted cultivars within a given crop. Within soybean, there is cultivar variation for light-saturated leaf photosynthesis (from 0.82 to 1.39 mg CO_2 m^{-2} s^{-1}), although the major variation falls within 0.92–1.17 mg CO_2 m^{-2} s^{-1}, with a mean of 1.05 mg CO_2 m^{-2} s^{-1} (Dornhoff and Shibles, 1970; Sinclair, 1980; Boote and Tollenaar, 1994). More recently released cultivars have higher light-saturated leaf photosynthesis than older cultivars that is associated primarily with increased SLW (Dornhoff and Shibles, 1970; Buttery et al., 1981; Wiebold et al., 1981). The Williams cultivar was somewhat exceptional in those studies, as it had low SLW and low leaf photosynthesis, but it was a good-yielding cultivar at that time. Most importantly, seed yield response to increasing leaf photosynthesis is less than proportional to the percentage increase in Pmax. Where the increase in leaf photosynthesis is not coupled to SLW change, seed yield is increased about 3–4% for each 10% increase in Pmax above the mid-point (Boote and Tollenaar, 1994), or about 5.1% in the simulations shown in Table 5. Similar results were reported by Kropff et al. (1995) for rice. When increase in Pmax is caused only by increased SLW, then the increase in seed yield is small (1–2%) for each 10% increase in Pmax because LAI is concurrently decreased (Boote and Tollenaar, 1994; also Table 5). Unfortunately, much of the reported variation in Pmax in soybean is associated with SLW. An exception is when prolonged leaf rate is allowed by higher leaf N concentration or by slower N mobilization (stay-green). When increase in Pmax is not coupled to SLW, seed yield response to increasing Pmax tends toward a saturating response. When Pmax is coupled to SLW, there is a clear optimum SLW and Pmax at which yield is maximum because the associated increased SLW results in lower LAI and lower light interception (see Fig. 20-3 of Boote and Tollenaar, 1994).

In our experience with different soybean cultivars, we have found only minimal differences in SLW of total canopy leaf area across soybean cultivars. So we feel this is not a particularly important trait. Environmental effects on SLW are more dominant and important than cultivar effects.

Table 5
Yield response to hypothetical variation in genetic traits of soybean simulated for 17 weather years for crops sown 2 May under rainfed conditions on a Nicollet soil at Ames, Iowa

Cultivar trait	Seed yield (kg ha^{-1})	Change (%)	Seed HI (frac.)	Node (No.)	Plant height (m)	Max. LAI (m^2 m^{-2})	Final biomass (kg ha^{-1})
Wm-82 std.	3160	–	0.447	20.3	1.08	5.96	7010
+ 10% higher leaf N	3254	3.0	0.454	20.3	1.08	5.81	7175
10% slower N mobil	3242	2.6	0.457	20.3	1.08	5.98	7087
+ 10% Pmax	3320	5.1	0.448	20.3	1.08	6.14	7348
+ 10% SLW	3196	1.1	0.444	20.4	1.07	5.68	7131
+ 10% R5–R7, same R7	3209	1.6	0.471	19.7	1.05	5.60	6734
+ 10% R5–R7, extend R7	3233	2.3	0.453	20.3	1.07	5.95	7085
Determinant type	3235	2.4	0.466	14.9	0.79	5.40	6843
Seed protein, 38.3–41.6%	3131	−0.9	0.450	20.3	1.07	5.95	6905
Seed oil, 19.2–22.8%	3057	−3.3	0.437	20.4	1.08	5.97	6942
Det, EFP[a]	3294	4.2	0.494	13.9	0.73	4.77	6577
Det, EFP, SLW	3330	5.4	0.494	13.9	0.73	4.46	6655
Det, EFP, Pmax	3475	10.0	0.495	13.9	0.74	4.99	6925

[a] Det, determinant cultivar; EFP, effective pod filling duration; SLW, specific leaf weight; Pmax, maximum leaf photosynthesis.

5.4. Effects of leaf N concentration and slower N mobilization

Although neither of these features are presently explicit cultivar traits in CROP-GRO, there is reported variation in leaf N concentration among soybean cultivars (Young et al., 1979; Sexton et al., 1998) and differences exist among peanut cultivars (Boote, unpublished). There is also evidence of slower N mobilization or stay-green in improved cultivars compared with older ones (Wells et al., 1982). Increase in leaf N concentration at a constant SLW will have the effect of increasing leaf photosynthesis. Initial biomass growth is less because of the added cost of assimilating extra N; however, the effect pays off later during mid-life cycle and during the reproductive period by increasing leaf photosynthesis as well as providing additional N to mobilize to seed. The modelled response to 10% increase in leaf N concentration for soybean (Table 5) is a 3.0% increase in seed yield and slight increase in seed HI (0.447 to 0.454). Even with no change in maximum leaf N concentration, yield increased 2.6% with 10% slower N mobilization which contributed to increase of HI from 0.447–0.457. There is a limit to yield gain with this feature because the mobilized C comes with the mobilized N and the cost of N assimilation is already paid for. Thus, it is best to mobilize most of the available N, but in a delayed manner.

5.5. Negative association of yield with percent protein or percent oil

Seed composition is known to vary among soybean cultivars, and plant breeders frequently report a negative association of yield with protein or oil concentration.

Being optimistic, the plant breeders hope to 'break' this 'genetic' linkage. With the CROPGRO-soybean model, we can easily increase protein or oil of the seed, with no changes to the other coefficients. Our simulations show that seed yield of soybean decreases 0.9% as protein increases from 38.3 to 41.6%, and decreases 3.3% as oil increases from 19.2 to 22.8% (Table 5). This response is attributed to additional energy cost required to synthesize protein and lipid as compared with cellulose-sugars-structure. In other words, this is a case of limited total assimilate. Previous simulations by Boote and Tollenaar (1994) showed that if breeders made soybeans with the composition of maize, they could increase soybean yield 30–40%. However, the energy and nutritional value of protein and lipids is higher, thus compensating the reduction in dry matter yield.

5.6. Additivity and interaction of traits with each other, and with management

For many traits illustrated in the simulations on Table 5, the trait effects appear to be primarily additive. From the example in Table 5, soybean yield was increased 1.6% from longer seed fill and 2.4% from determinant trait, whereas the combination of the two traits gave 4.2% response (Table 5) which is primarily additive. Including a third trait of increased photosynthesis gave a slightly more than additive 10% yield response (sum of the parts was $1.6 + 2.4 + 5.1 = 9.1\%$). However, Boote and Tollenaar (1994) found interaction of some traits with crop management. In their simulations, the determinate trait was not an advantage at wide row spacing and low plant population (where yield was not improved because of losses in light interception), but it was advantageous when placed in narrow row spacing and high plant population. Similarly, the determinate trait was more advantageous in a high leaf photosynthesis background.

6. Crop modelling to determine ideotypes to design plants for target environments

An example of an improved plant type design that had a major impact on agricultural production is the short stature design for cereals that was developed in the 1960s. These early breeding efforts had their focus on wide adaptation. Although the wide adaptation concept worked for some systems like the homogeneous irrigated rice systems in the tropics and subtropics and wheat in South Asia, it has been recognized that strong genotype×environment×management interactions in most agricultural systems make it necessary to use genotypes that are adapted to the specific agro-environments. To meet the challenges to increase food production, ways have to be found to improve the productivity and profitability of agricultural production systems. That can only be done by using a systems approach that helps to optimize local agricultural production systems. That will require not only improved crop management systems, but also new crop types for different agro-environments.

Hammer et al. (1999) proposed that crop models have the potential to interpret and predict performance of individual genotypes in different managements and

environments. They also concluded that crop models based on physiologically sound mechanisms, can quantify and integrate crop responses to genetic, environmental, and management factors, and therefore, can be used as a tool to study genotype by environment interactions and genotypes suited for particular climatic zones.

To analyze the need for different plant types for different environments, Kropff et al. (1995) determined the opportunities for using models to optimize the performance of specific genotypes in target environments with rice and potato as case studies. Differences in the requirements for rice varieties for the different seasons became clear in the simulations for the tropical wet season and the dry season. However, the basic traits were similar. Model analyses showed that the key to the optimization of potential yield in potato was the proper timing of the moment of tuber initiation. If that moment is too late for a particular environment, too much foliar tissue is formed and the tuber growth period can be too short. If tuber initiation is too early, not enough foliar tissue is formed to sustain growth until the end of the growing season. LINTUL-POTATO was used to determine the cultivar characteristics with respect to temperature and daylength response that give the highest yield in a particular environment. This means that for each environment the optimal moment of tuber initiation can be calculated and only genotypes with such desired earliness have to be evaluated in this environment.

Kropff et al. (1995) and in more detail, Aggarwal et al. (1997) determined the importance of various traits for maximum yield potential in the tropical dry season and the wet season using ORYZA1 (Kropff et al., 1994) which is based on the Wageningen models SUCROS (Spitters et al., 1989) and MACROS (Penning de Vries et al., 1989). Crop characteristics of IR72 and weather data of Los Baños for 1992 were used. The critical model parameter values were varied with respect to those of IR72 to simulate the effect of a change in specific leaf area, spikelet growth factor, potential grain weight, maximum leaf N concentration, and crop development rates during juvenile phase and grain filling period. The study was conducted at three levels of N management: normal practice in breeders plots, recommended N for high yield potential and growth-rate driven N management. Results showed that no trait individually or in combination could provide more than 5% yield increase in the normal N management practice. These genotypes were not able to express these traits in these somewhat limited-N environments. With improved N management, yield potential was predicted to be significantly increased ($>30\%$) by an increased sink capacity, maintenance of high leaf N content and a longer grain filling duration (Aggarwal et al., 1997).

Differential drought patterns are a major cause for reported genotype by environment interactions. Chapman et al. (2000) used a sorghum simulation model for environmental characterization of long-term historical weather data for different regions. They classified several different drought patterns that led to different growth and yield patterns. Cultivar selection under these different environments could lead to selection of different cultivar types. Even more important is that the occurrence of a given drought pattern from one year to the next, relative to variety trials conducted over a 3-year span (of good or of poor weather) could lead to selection of a variety that will not be the best over the long term. So one valuable role of crop

models is to evaluate optimum cultivar types over long-term weather, rather than the short-term 1–3 year duration typical prior to release of a new cultivar.

Using crop models as tools, Bindraban (1997) showed that at optimum planting dates and optimum management conditions in a highly productive site (Obregon, north-west Mexico) the photosynthesis capacity of the canopy exceeded the actual yield. Residual reserves in the stem were higher in crops grown under these optimal conditions than in low-yielding crops grown with insufficient nitrogen. Optimally managed crops that were sown later, and therefore experienced higher temperatures, gave lower yields as the rapidly declining post-anthesis leaf area strongly reduced photosynthesis capacity. Bindraban (1997) concluded that the sink rather than the source limits yield potential in highly productive sites, while the reverse is true for crops at lower yielding levels. Yield potential under optimum managed conditions in productive sites should therefore be improved by increasing sink capacity, either through increasing kernel number or kernel weight. With the aid of crop modelling it was shown that different crop characteristics ought to be adjusted depending on environmental conditions. Under the optimum planting conditions of Obregon, kernel number could be increased by increasing crop growth rate during the period between early booting and final anthesis. For late planted crops in Obregon, an elongation of this period appeared more appropriate (Bindraban et al., 1998).

While solving for genetic coefficients and site characteristics (soil water holding traits and fertility) from soybean variety trials across differing environments (described previously in this paper), we found that the crop model successfully reproduced part of the observed G×E interaction for different simulated cultivar pairs (Figs. 8 and 9). The observed steeper slope of Thomas (lower yield in poor
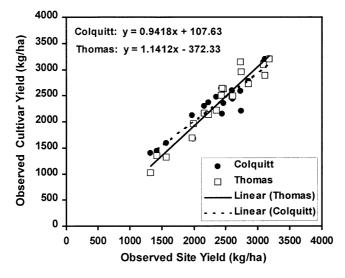

Fig. 8. Observed cultivar yield of Colquitt and Thomas versus observed site yield, taken from Georgia variety trial data (Mavromatis et al., 2001). The respective regression lines are shown.
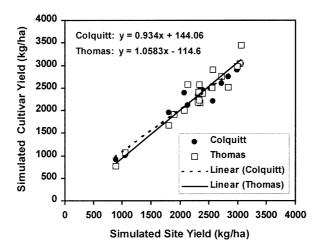
Fig. 9. Simulated cultivar yield of Colquitt and Thomas versus simulated site yield, after solving cultivar and site traits from the Georgia variety trial data (Mavromatis et al., 2001). The respective regression lines are shown.

environments and higher yield in good environments) compared with Colquitt (Fig. 8) was substantially captured by the model (Fig. 9) after solving cultivar traits (slopes of Thomas and Colquitt differed significantly, both in observed and simulated pair comparisons). We believe that some of the G×E in yield is related to different phenologies which the optimizer solves quite efficiently. In addition, G×E for yield also is influenced by soil fertility and soil water availability (solved for sites), as well as genetic traits influencing yield through photosynthesis and seed HI. Other reasons for G×E for yield include effects of excess water and pests, which most models do not presently account for, but could address in the future. It is encouraging that G×E interactions noted here resulted naturally from the process of solving for genotypic traits across many environments, rather than the need to explicitly 'hard-code' such G×E features as genetic coefficients. We view Fig. 6 as another example of how such G×E interactions can occur, where traits of longer grain filling period or slower N mobilization can enhance yield considerably in good environments, with much less or no effect in drought-prone years.

Analyses using crop models can improve our insight into the quantitative impact of crop characteristics on yield potential. With this information, plant breeders can concentrate their efforts on the quantitatively most important traits that need to be considered for improving yields, while modellers can enhance their crop models to better mimic such genetic differences.

7. The future — expanded and extended physiological coefficients

Crop modellers are well aware that their models inadequately describe differential cultivar tolerances for drought, cold, heat, diseases, and insect pests. Thus, there is

potential for an expanded set of physiological coefficients that describe these various tolerances. Even where there is motivation to address some of these issues, the problem is time, resources, and availability of detailed knowledge. Help may be on the way, in the form of molecular genetics, if the molecular DNA knowledge of released cultivars is adequately measured and made available, and if linkages and responses of those genetics to given pests, cold, heat, and drought are described. Hammer et al. (1999) proposed that crop models could be used to enhance our understanding of molecular genetic technologies, by characterizing those climatic and management situations when a given quantitative trait loci (QTL) is associated with yield and when the QTL is not associated with yield. They gave a hypothetical example of increased yield associations with early flowering (allowing longer seed fill duration), and suggested that this linkage would be broken under terminal drought stress, in which case the early flowering cultivar is best because of drought escape. The crop models can give us logical reasons for this effect on yield ranking. Likewise, the study by Mavromatis et al. (2001) showed that the crop model could predict G×E and rank reversals associated with different field environments. Yin et al. (1999a, 1999b and unpublished) conducted studies with barley where they simulated yield differences in barley cultivars using a barley model based on SUCROS and ORYZA1. They genetically identified important traits used by the model using QTL techniques. For example, they identified one to three QTLs for specific leaf area (SLA) at different measurement times in the season. They demonstrated that model prediction based on measured traits of the inbred lines was as accurate as model predictions based on QTL predicted model traits.

There is evidence to support the feasibility of a gene-based approach for modelling cultivar differences in crop simulation models, even using present knowledge of known genes (White and Hoogenboom, 1996; Hoogenboom et al., 1997). White and Hoogenboom (1996) devised a gene-based simulation model, GeneGRO, that integrates the action of seven genes for *Phaseolus vulgaris*. The specific genes affect photoperiod sensitivity, development rate prior to flowering, development rate after flowering, growth habit, pod and seed growth. In an independent test over 39 cultivars in 14 trials, the GeneGRO model accounted for 75% of the variation in time to flower, 68% for maturity, 39% for seed size (mass), but only 11% for seed yield (Hoogenboom et al., 1997). The authors thought it remarkable that this much variation in phenology could be explained by as few as four genes, and discussed minor improvements to consider additional gene effects on phenology. They (and we) are not surprised by the low variation in seed yield accounted for, because seed yield integrates all environmental and site effects, and the simulations (and observed data) would also be influenced by errors in model assumptions, incorrect site-specific inputs, and experimental errors. GeneGRO considers only a few of the many 'yield' genes beyond phenology, and the authors discussed examples of known and yet-to-be-discovered quantitative genes that may affect yield. We and these authors conclude that knowledge of genetic information obtained by molecular genetic techniques such as QTLs should facilitate gene-based crop simulation models, provided further information is obtained on physiological mechanisms of action of these genes or genetic sequences.

An extended set of genetic coefficients could be added to include differential pest-resistances and differential drought-tolerance of cultivars. There is good evidence of cultivar differences in pest resistance that could be assigned as genetic traits; however, the larger problem is how to predict pest occurrence or the need to input pest pressures for each new field and season (as was done by Batchelor et al., 1993). For drought-tolerance, the literature is full of hypothetical cases, but there is limited real evidence on which to assign differential mechanisms of drought-tolerance to cultivars. E.L. Piper concluded after an exhaustive literature search (unpublished), that there was little hard evidence of the mechanisms of different soybean cultivar tolerance to drought, except for deeper rooting, possibly allowed by differential tolerance of acid or high aluminium subsoils. Piper used model simulation approaches to evaluate several hypothetical mechanisms of cultivar differences that could lead to differential drought tolerance (ability to yield under droughts).

We conclude there is excellent potential for better integration of crop modelling and breeding programmes, where crop models can be used to evaluate past genetic improvement, to hypothesize new genotypes, to propose ideotypes for target environments, and to solve for genetic coefficients from variety trials. Crop model links to extensive international (and national) data bases such as ICIS (International Crop Information System), as described by Hunt et al. (2001), can provide access to information and ability to routinely derive genetic coefficients. Once available, these genetic coefficients will allow an increased level of crop model applications for management, policy, etc. Future interactions with molecular geneticists offers yet another avenue for both deriving genetic coefficients (based more on the genetics than on performance trials), and for assisting breeders in determining the responses of new genetic materials in field environments.

References

Aggarwal, P.K., Kropff, M.J., Cassman, K.G., Ten Berge, H.F.M., 1997. Simulating genotypic strategies for increasing rice yield potential in irrigated, tropical environments. Field Crops Research 51, 5–17.

Batchelor, W.D., Jones, J.W., Boote, K.J., Pinnschmidt, H.O., 1993. Extending the use of crop models to study pest damage. Trans. ASAE 36, 551–558.

Bindraban, P.S., 1997. Bridging the Gap Between Plant Physiology and Breeding. Identifying Traits to Increase Wheat Yield Potential Using Systems Approaches. PhD thesis, Wageningen Agricultural University, Wageningen, The Netherlands.

Bindraban, P.S., 1999. Impact of canopy nitrogen profile in wheat on growth. Field Crops Research 63, 63–77.

Bindraban, P.S., Sayre, K.D., Solis-Moya, E., 1998. Identifying factors that determine kernel number in wheat. Field Crops Research 58, 223–234.

Boote, K.J., Jones, J.W., 1986. Applications of, and limitations to, crop growth simulation models to fit crops and cropping systems to semi-arid environments. In: Bidinger, F.R., Johansen, C. (Eds.). Drought research priorities for the dryland tropics. International Crops Research Institute for the Semi-Arid Tropics, Pantancheru, A.P. 502 324, India, pp. 63–75.

Boote, K.J., Tollenaar, M., 1994. Modeling genetic yield potential. In: Boote, K.J., Bennett, J.M., Sinclair, T.R., Paulsen, G.M. (Eds.). Physiology and Determination of Crop Yield. ASA-CSSA-SSSA, Madison, WI, USA, pp. 533–565.

Boote, K.J., Jones, J.W., 1998. Crop modeling and the Soybean Seed Industry. Proceedings, American Seed Trade Association. Washington, DC, USA.

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

Buttery, B.R., Buzzell, R.I., Findlay, W.I., 1981. Relationships among photosynthetic rate, bean yield and other characters in field-grown cultivars of soybean. Can. J. Plant Sci. 61, 191–198.

Chapman, S.C., Cooper, M., Hammer, G.L., Butler, D., 2000. Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. Aust. J. Agric. Res. 51, 209–221.

De Wit, C.T., Penning de Vries, F.W.T., 1983. Crop growth models without hormones. Neth. J. Agric. Sci. 31, 313–323.

Dornhoff, G.M., Shibles, R.M., 1970. Varietal differences in net photosynthesis of soybean leaves. Crop Sci. 10, 42–45.

Duncan, W.G., McCloud, D.E., McGraw, R.L., Boote, K.J., 1978. Physiological aspects of peanut yield improvement. Crop Sci. 18, 1015–1020.

Dunphy, E.J., Hanway, J.J., Green, D.E., 1979. Soybean yields in relation to days between specific developmental stages. Agron. J. 71, 917–920.

Egli, D.B., Orf, J.H., Pfeiffer, T.W., 1984. Genotypic variation for duration of seedfill in soybean. Crop Sci. 24, 587–592.

Elwell, D.L., Curry, R.B., Keener, M.E., 1987. Determination of potential yield-limiting factors of soybeans using SOYMOD/OARDC. Agric. Systems 24, 221–242.

Goudriaan, J., Van Laar, H.H., 1994. Modelling Potential Crop Production Processes. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Grimm, S.S., Jones, J.W., Boote, K.J., Hesketh, J.D., 1993. Parameter estimation for predicting flowering date of soybean cultivars. Crop Sci. 33, 137–144.

Grimm, S.S., Jones, J.W., Boote, K.J., Herzog, D.C., 1994. Modeling the occurrence of reproductive stages after flowering for four soybean cultivars. Agron. J. 86, 31–38.

Hammer, G.L., 1998. Crop modelling: current status and opportunities to advance. Acta Hort. 456, 27–36.

Hammer, G.L., Butler, D., Muchow, R.C., Meinke, H., 1996. Integrating physiological understanding and plant breeding via crop modelling and optimisation. In: Cooper, M., Hammer, G.L. (Eds.), Plant Adaptation and Crop Improvement. CAB International, Wallingford, UK, pp. 419–441.

Hammer, G.L., Chapman, S.C., Snell, P., 1999. Crop simulation modelling to improve selection efficiency in plant breeding programs. Proceedings of the ninth Assembly Wheat Breeding Society of Australia, Toowoomba, September 1999, pp. 79–85.

Hoogenboom, G., Jones, J.W., Wilkens, P.W., Batchelor, W.D., Bowen, W.T., Hunt, L.A., Pickering, N.B., Singh, U., Godwin, D.C., Baer, B., Boote, K.J., Ritchie, J.T., White, J.W., 1994. Crop models. In: Tsuji, G.Y., Uehara, G., Balas, S. (Eds.), DSSAT Version 3, Vol. 2-2. University of Hawaii, Honolulu, HI, pp. 95–244.

Hoogenboom, G., White, J.W., Acosta-Gallegos, J., Gaudiel, R.G., Myers, J.R., Silbernagel, M.J., 1997. Evaluation of a crop simulation model that incorporates gene action. Agronomy J. 89, 613–620.

Hunt, L.A., Pararjasingham, S., Jones, J.W., Hoogenboom, G., Imamura, D.T., Ogoshi, R.M., 1993. GENCALC: software to facilitate the use of crop models for analyzing field experiments. Agron. J. 85, 1090–1094.

Hunt, L.A., White, J.W., Hoogenboom, G., 2001. Agronomic data: advances in documentation and protocols for exchange and use. Agric. Systems 70 (2–3), 477–492.

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

Jackson, P., Robertson, M., Copper, M., Hammer, G.L., 1996. The role of physiological understanding in plant breeding: from a breeding perspective. Field Crops Research 49, 11–37.

Kropff, M.J., Van Laar, H.H. (Eds.), 1993. Modeling Crop–Weed Interactions. CAB International, Wallingford, Oxon, UK. International Rice Research Institute, Los Baños, Philippines.

Kropff, M.J., Van Laar, H.H., Matthews, R.B. (Eds.), 1994. ORYZA1: an ecophysiological model for

irrigated rice production. SARP Research Proceedings, AB-DLO. Wageningen and IRRI, Los Baños, Philippines.

Kropff, M.J., Haverkort, A.J., Aggarwal, P.K., Kooman, P.L., 1995. Using systems approaches to design and evaluate ideotypes for specific environments. In: Bouma, J., Kuyvenhoven, A., Bouman, B.A.M., Luyten, J.C., Zandstra, H.G. (Eds.), Eco-Regional Approaches for Sustainable Land Use and Food Production. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 417–435.

Landivar, J.A., Baker, D.N., Jenkins, J.N., 1983a. Application of GOSSYM to genetic feasibility studies. I. Analyses of fruit abscission and yield in okra-leaf cottons. Crop Sci. 23, 497–504.

Landivar, J.A., Baker, D.N., Jenkins, J.N., 1983b. Application of GOSSYM to genetic feasibility studies. II. Analyses of increasing photosynthesis, specific leaf weight and longevity of leaves in cotton. Crop Sci. 23, 504–510.

Mavromatis, T., Boote, K.J., Jones, J.W., Irmak, A., Shinde, D., Hoogenboom, G., 2001. Developing genetic coefficients for crop simulation models with data from crop performance trials. Crop Sci. 41, 40–51.

Penning de Vries, F.W.T., Jansen, D.M., Ten Berge, H.F.M, Bakema, A., 1989. Simulation of Ecophysiological Processes of Growth of Several Annual Crops. Simulation Monographs 29. Pudoc. Wageningen and IRRI, Los Baños, Philippines.

Piper, E.L., Boote, K.J., Jones, J.W., Grimm, S.S., 1996. Comparison of two phenology models for predicting flowering and maturity date of soybean. Crop Sci. 36, 1606–1614.

Sexton, P.J., Batchelor, W.D., Boote, K.J., Shibles, R., 1998. Evaluation of CROPGRO for prediction of soybean nitrogen balance in a midwestern environment. Trans. ASAE 41, 1543–1548.

Shorter, R., Lawn, R.J., Hammer, G.L., 1991. Improving genotypic adaptation in crops: a role for breeders, physiologists and modellers. Expl. Agric. 27, 155–175.

Sinclair, T.R., 1980. Leaf CER from post-flowering to senescence of field-grown soybean cultivars. Crop Sci. 20, 196–200.

Spitters, C.J.T., Van Keulen, H., Van Kraalingen, D.W.G., 1989. A simple and universal crop growth simulator: SUCROS87. In: Rabbinge, R., Ward, S.A., Van Laar, H.H. (Eds.), Simulation and Systems Management in Crop Protection. Simulation Monographs, Pudoc, Wageningen, The Netherlands, pp. 147–181.

Swank, J.D., Egli, D.B., Pfeiffer, T.W., 1987. Seed growth characteristics of soybean genotypes differing in duration of seed fill. Crop Sci. 27, 85–89.

Wells, R., Schulze, L.L., Ashley, D.A., Boerma, H.R., Brown, R.H., 1982. Cultivar differences in canopy apparent photosynthesis and their relationship to seed yield in soybeans. Crop Sci. 22, 886–890.

Whisler, F.D., Acock, B., Baker, D.N., Fye, R.E., Hodges, H.F., Lambert, J.R., Lemmon, H.E., McKinion, J.M., Reddy, V.R., 1986. Crop simulation models in agronomic systems. Adv. in Agron. 40, 141–208.

White, J.W., Hoogenboom, G., 1996. Simulating effects of genes for physiological traits in a process-oriented crop model. Agron. J. 88, 416–422.

Wiebold, W.J., Shibles, R., Green, D.E., 1981. Selection for apparent photosynthesis and related leaf traits in early generations of soybeans. Crop Sci. 21, 969–973.

Wilkerson, G.G., Jones, J.W., Boote, K.J., Ingram, K.T., Mishoe, J.W., 1983. Modeling soybean growth for crop management. Trans. of ASAE 26, 63–73.

Yin, X., Kropff, M.J., Stam, P., 1999a. The role of ecophysiological models in QTL analysis: the example of specific leaf area in barley. Heredity 82, 415–421.

Yin, X., Stam, P., Dourleijn, J.C., Kropff, M.J., 1999b. AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. Theoretical Applied Genetics 99, 244–253.

Yin, X., Kropff, M.J., Aggarwal, P.K., Peng, S., Horie, T., 1997. Optimal preflowering phenology of irrigated rice for high yield potential in three Asian environments: a simulation study. Field Crops Res. 51, 19–27.

Young, J.K., Whisler, F.D., Hodges, H.F., 1979. Soybean leaf N as influenced by seedbed preparation methods and stages of growth. Agron. J. 71, 568–573.