GPFARM Plant Model Parameters: Complications of Varieties and the Genotype × Environment Interaction in Wheat

G. S. McMaster, J. C. Ascough II, M. J. Shaffer, L. A. Deer–Ascough, P. F. Byrne,

D. C. Nielsen, S. D. Haley, A. A. Andales, G. H. Dunn

Abstract. *The USDA–ARS Great Plains Framework for Agricultural Resource Management (GPFARM) decision support system was developed to assist Great Plains producers in making economically viable and environmentally sound strategic plans for whole farm and ranch systems. A major user requirement for GPFARM is to supply the default plant parameters required to simulate crop growth. Developing this plant parameter database is difficult because varietal differences, caused by a genotype by environment (G × E) interaction, increases parameter uncertainty and variability. This article examines species–based plant parameter sets for simulating winter wheat (*Triticum aestivum *L.) yield responses, explores the significance of the G × E interaction on simulating varietal grain yield, and investigates whether simple adjustments to a species–based plant parameter database can improve simulation of varietal differences across environments. Three plant parameter sets were evaluated against observed yield data for six locations in eastern Colorado: (1) the Default parameter set used best estimates from EPIC–based plant parameter databases, (2) the Dryland Agroecosystems Project (DAP) parameter set further calibrated the default plant parameters against observed yield data for Colorado, and (3) the Theory parameter set modified DAP parameters based on whether irrigated or dryland conditions were simulated. The Theory parameter set simulated yield the best when pooling varieties over environments and locations. However, no parameter set could simulate all the different varietal yield responses to environmental conditions (irrigated or dryland) due to the diverse G × E interactions. The Theory parameter set best simulated the wheat variety TAM 107 across diverse locations, with little bias for either irrigated or dryland conditions. Simple adjustments to a few plant parameters based on whether dryland or irrigated conditions were simulated improved the species–based plant parameter approach used in GPFARM. However, until a better mechanistic representation of the G × E interaction is incorporated into existing plant growth models, opportunities for improving yield response to environmental conditions and management will be limited.*

Keywords. *Crop growth models, Crop yield, Genotype by environment interaction, Plant parameters, Simulation models.*

Agricultural software developers are increasingly delivering products (e.g., decision support systems; simulation models; and budgeting, record keeping, and irrigation/nitrogen/weed control management programs or tools) for use by farmers and ranchers. Unfortunately, most agricultural software is rarely adopted or used on the farm or ranch, especially decision support systems (DSS) and simulation models (Ascough et al., 1999, 2002b). Perhaps the most important reasons for not adopting these products are because they are viewed as too difficult to use and the investment of time and effort to learn, set up, and run the software is not returned in value to the producer.

The USDA–ARS Great Plains Systems Research Unit has developed a decision support system named GPFARM (Great Plains Framework for Agricultural Resource Management). GPFARM 2.5 encompasses stand–alone components such as a user interface, simulation model, and databases (Ascough et al., 2002a; McMaster et al., 2002b; Shaffer et al., 2000) that, when used in conjunction with other components (e.g., farm economic budgeting and multicriteria decision analysis modules), provides a unique decision support tool for farmers and ranchers. GPFARM was designed to provide decision support for use by farmers and ranchers in strategically managing their complex agricultural systems. From the inception of the GPFARM project, producer input on system characteristics and functionality has been solicited. The crop growth component that simulates growth and yield also had to meet these requirements, and still be sufficiently robust to cover the diverse environments, cropping systems, and management practices employed across the Great Plains. Providing a robust crop growth model for producers necessitated that parameters for simulating growth be provided in a

Article was submitted for review in October 2002; approved for publication by the Soil & Water Division of ASAE in May 2003.

The authors are **Gregory S. McMaster,** Agronomist, **James C. Ascough II, ASAE Member Engineer,** Hydraulic Engineer, **Marv J. Shaffer,** Soil Scientist, **Lois A. Deer–Ascough,** Agricultural Engineer, **Allan A. Andales, ASAE Member Engineer,** Visiting Scientist, and **Gale H. Dunn,** Soil Scientist, USDA–ARS Great Plains Systems Research Unit, Fort Collins Colorado; **David C. Nielsen,** Agronomist, USDA–ARS Central Great Plains Research Station, Akron, Colorado; and **Pat F. Byrne,** Associate Professor, and **Scott D. Haley,** Associate Professor, Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado. **Corresponding author:** Dr. Gregory S. McMaster, USDA–ARS, GPSR, 301 S. Howes St., Room 353, Fort Collins, CO 80521; phone: 970–490–8340; fax: 970–490–8310; e–mail: greg.mcmaster@ars.usda.gov.

Williams et al., 1984, 1989). It has been further modified in GPFARM and incorporates some elements from the Agricultural Land Management Alternatives with Numerical Assessment Criteria (ALMANAC) model (Kiniry et al., 1992). A single model is used for simulating multiple crops by changing model parameters. Stress factors for water and nitrogen are computed using inputs from other independent models within GPFARM.

The crop growth component can be characterized as using the energy– or carbon–driven approach common in plant growth modeling. Potential daily biomass accumulation is based on the interception of light by the canopy (as represented by the LAI and light extinction coefficients) and an energy–to–biomass conversion factor. Limiting abiotic resources are reflected in growth constraint factors (temperature, water, and N) reducing the potential daily biomass accumulation. Carbon and N are partitioned to plant components (e.g., leaves, roots, grain) based on phenological growth stage.

Phenological development of the crop is based on thermal time using daily heat unit accumulation. Daily heat units are computed using the equation:

$$HU_i = \frac{(T_{\max,i} + T_{\min,i})}{2} - T_{base,j} \qquad (1)$$

where
- HU = heat units (°C) on day i
- $T_{\max}$ = maximum temperature (°C) on day i
- $T_{\min}$ = minimum temperature (°C) on day i
- T_{base} = crop–specific base temperature (°C) of crop j.

No growth occurs at or below T_{base}, and there is no upper temperature limit.

A heat unit index (HUI) ranging from 0 at planting to 1 at physiological maturity is computed as follows:

$$HUI_i = \frac{\sum_{k=1}^{i} HU_k}{PHU_j} \qquad (2)$$

where
- HUI = heat index for day i
- k = counter representing the summation of days
- PHU = potential heat units required to reach maturity for crop j.

Several equations are used in determining daily potential biomass production. Interception of photosynthetic active radiation (PAR) is estimated with Beer's law (Monsi and Saeki, 1953):

$$PAR_i = 0.02092(RA_i)(1.0 - e^{-0.65LAI})_i \qquad (3)$$

where
- PAR = photosynthetic active radiation (MJ m^{-2})
- RA = solar radiation (Langleys)
- LAI = leaf area index
- i = day of the year.

Potential biomass production per day is estimated with the equation (Montieth, 1977):

$$\Delta BP_i = 0.0001(BE_j)(PAR_i) \qquad (4)$$

where
- ΔBP = potential increase in total biomass on day i (kg m^{-2})
- BE = energy to biomass conversion parameter for crop j (kg MJ^{-1}).

Actual daily biomass accumulation is determined by Leibig's Law of the Minimum. The daily potential biomass accumulation (eq. 4) is adjusted daily if one of the plant stress factors (water, N, or temperature) is less than 1.0 using the equation:

$$\Delta B_i = (\Delta BP_i)(REG) \qquad (5)$$

where REG is the crop growth regulating factor (the minimum of the water, N, and temperature stress factors) calculated for day i. The adjusted daily total biomass production (ΔB_i) is accumulated through the growing season.

The water stress factor is computed by considering supply and demand in the equation:

$$WS_i = \frac{\sum_{l=1}^{nl} u_{l,i}}{E_{p,i}} \qquad (6)$$

where
- WS = water stress factor (0–1)
- u_l = plant water use in soil layer l (mm)
- nl = number of soil layers
- E_p = potential plant transpiration (mm)
- i = day of the year.

The N stress factor is computed by considering the N demand for biomass production and amount of plant N uptake in the equation:

$$NS_i = \frac{\sum_{l=1}^{nl} V_{l,i}}{N_{p,i}} \qquad (7)$$

where
- NS = N stress factor (0–1)
- V_l = plant N (NO_3–N + NH_4–N) uptake in soil layer l (kg/ha)
- N_p = plant N demand
- i = day of the year.

N_p is calculated as a percentage of daily total biomass production and varies depending on crop growth stage based on plant parameters BN1, BN2, and BN3 (table 1) for emergence, mid–season, and maturity, respectively.

The temperature stress factor is computed with the equation:

$$TS_i = \sin\left(\frac{\pi}{2}\,\frac{T_{ave,i} - T_{base,j}}{T_{opt,j} - T_{base,j}}\right) \qquad (8)$$

where
- TS = temperature stress factor (0–1)
- T_{ave} = average daily temperature (°C)
- T_{opt} = optimum temperature (°C) for crop j
- i = day of the year.

Crop yield for annual crops is estimated using the harvest index concept, which is adjusted throughout the growing season according to water stress constraints:

$$YLD_j = (HIA_j)(B_{AG}) \qquad (9)$$

where
- YLD = crop yield (kg m^{-2}) at harvest for crop j
- HIA = adjusted harvest index for crop j

(e.g., energy to biomass conversion, harvest index, maximum LAI, thermal time from planting to maturity), found via experience in using the model, were modified to determine if the yield response improved compared to county yield averages. The resulting final default parameter set is denoted as Default in this article; table 1 lists values for winter wheat supplied to GPFARM 2.5 users.

The Default parameter set for winter wheat was further examined and run on a subset of location–treatment–years for the Dryland Agroecosytems Project (DAP) discussed in the Evaluation Data Sets section below (Peterson et al., 2001). Only the wheat–fallow rotation for the summit position at each location was used in calibration (a total of 27 out of 94 location–treatment–years). As in creating the Default plant parameter database, important parameters influencing yield were informally adjusted based on existing data (e.g., McMaster, 1997; McMaster and Smika, 1988; McMaster et al., 1992, 1994) for the variety TAM 107 and "expert opinion" for winter wheat in the western Central Great Plains until simulated yield was improved (Andales et al., 2003). The parameters adjusted were harvest index (HI), maximum potential LAI (XMXLAI), heat unit index when leaf area index begins to decline (DLAI), thermal time from sowing to emergence and maturity (CRIT, GDDMAX), base (BTEMP) and optimum (OTEMP) temperature, and maximum rooting depth (RDMAX). This parameter set is denoted DAP in this article, and values are listed in table 1.

The Theory parameter set (table 1) was developed to address two problems of using species–based plant parameter databases such as the Default and DAP parameter sets. The first problem is that producers often pick varieties they perceive as best adapted for the type of farming they practice, such as irrigated or dryland. Varieties selected for irrigated or dryland conditions can differ greatly in traits, and thus parameter values should likely be changed, but varietal differences are not typically included in plant parameter databases. The second problem relates to simulation of certain processes such as the thermal time from sowing to emergence and maturity. The GPFARM crop growth model assumes a static parameter value (GDDMAX) for the thermal time for the species (or variety), yet it is undeniable that this value should respond to environmental conditions (particularly water stress) other than merely temperature (McMaster, 1997). Therefore, we theorized that simulating grain yield responses might be improved by having different parameters based on dryland and irrigated conditions, which would be an indirect approach to incorporating varietal differences that is a simple refinement easily adapted for all crops. The following parameters were modified for irrigated conditions only: GDDMAX (increased), XMXLAI (increased), DLAI (delayed), BEINP (increased), and RDMAX (decreased). The rationale for direction of change in the parameters is based on fundamental physiological principles. For instance, the onset of leaf senescence (DLAI) and maximum LAI (XMXLAI) is clearly influenced by water availability (McMaster et al., 1992), and this is not accounted for in the model.

Evaluation Data Sets

The primary data used to evaluate the winter wheat yield predictions were based on a two–year study conducted at the Colorado State University Agricultural Research Development and Education Center (denoted ARDEC; 40° 39′ N, 105° 00′ W, 1534 m elevation; fine, smectitic, mesic, Aridic Arguistoll) and the USDA–ARS Central Great Plains Research Station in Akron, Colorado (denoted Akron; 40° 09′ N, 103° 09′ W, 1384 m elevation; fine, smectitic, mesic, Pachic Arguistoll), both initiated in the fall of 1999. Twelve winter wheat varieties (or 10 varieties for year 1 at Akron; listed in fig. 1) differing in presumed heat and drought tolerance were grown under dryland and irrigated conditions. Most varieties are commonly used in this region, but several are adapted to other environments (e.g., Norstar and Siouxland). The experimental design was a split plot with dryland/irrigated conditions the main plot factor and variety the subplot. Replications differed with locations and year: two replications for 1999–2000 at both locations, and three (Akron) and four (ARDEC) replications for 2000–2001. Nitrogen fertilizer was applied at planting to meet recommended levels based on soil tests prior to planting.

Other data were obtained from the long–term DAP study (Peterson et al., 2001) for three locations in eastern Colorado (near Sterling, Stratton, and Walsh) initiated in 1986. Wheat–fallow, wheat–corn–fallow, and wheat–corn–millet–fallow rotations were grown under no–tillage management at different topographic positions of a catena. The winter wheat variety TAM 107 was used through 1998. Nitrogen and phosphorus fertilizer was applied at planting to meet recommended levels based on soil tests prior to planting.

The final data set was from a 6–year study conducted at the Colorado State University Horticultural Farm in Fort Collins, Colorado (denoted Hortfarm; 40° 36′ N, 104° 59′ W, 1515 m elevation) initiated in 1992 on a Nunn clay loam soil (fine, smectitic, mesic, Aridic Arguistoll; McMaster et al., 2002a). A split–plot design, with tillage being the main plot and residue rate being the subplot, with four replications was used with the cultivar TAM 107 in a wheat–fallow cropping system. Soil tests prior to planting indicated that only in the last year was N and P fertilizer (38 kg N ha^{-1} as 32% urea ammonium nitrate solution, and 9.5 kg P ha^{-1} as liquid ammonium polyphosphate: 10–34–0) required to meet recommended levels.

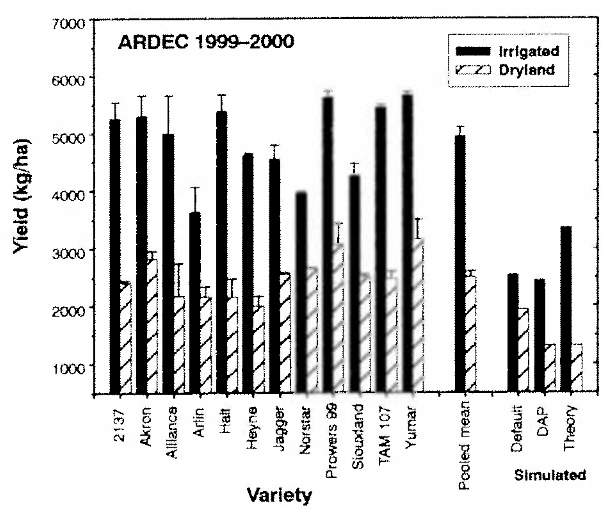

Figure 1. Observed grain yield (with 1 SE bar) for 1999–2000 at ARDEC, Colorado, for dryland and irrigated treatments compared to simulated grain yield for different parameter data sets. Simulated values are the three rightmost sets of bars.

Table 2. Statistical results of simulating grain yield (kg/ha) at ARDEC and Akron locations using the three plant parameter data sets (table 1). Observed yield was calculated using the mean of 12 cultivars for a site–year treatment. Simulation results are presented in figures 1–4.

Statistical Measure[a]	Default	DAP	Theory
Irrigated and dryland treatments			
RMSE	1158	1464	2761
SRES	6361	8821	4313
SARES	7297	9523	6430
Paired t–test (P)	0.04	0.02	0.12
Dryland treatments			
RMSE	622	889	889
SRES	1152	2136	2136
SARES	2089	2838	2838
Paired t–test (P)	0.43	0.28	0.28
Irrigated treatments			
RMSE	1514	1870	1056
SRES	5208	6685	2177
SARES	5208	6685	3592
Paired t–test (P)	0.06	0.04	0.37

[a] RMSE = root mean square error.
SRES = sum of residuals.
SARES = sum of absolute residuals.
P = probability that there is no difference between observed and simulated yield.

parameters is required to capture the varietal yield response. In our case, if the individual varieties are compared to the simulated yield for the three parameter sets, we would expect to see a consistent bias if the particular variety did not match simulated yields for the parameters chosen. This was not found statistically (data not shown), nor is it apparent in figures 1–4.

To further explore the possible importance of a G × E interaction on yield, we examined the percent yield loss [(irrigated – dryland yield)/irrigated yield × 100] rankings to water stress treatments between years at ARDEC and Akron (figs. 5 and 6). Except for two varieties at each location (Norstar and Arlin at ARDEC, Norstar and Prowers 99 for Akron), greater yield loss was observed in 2000 than in 2001. This was expected, as precipitation was less in 2000 than in 2001 (101 mm and 111 mm less precipitation from September through July for ARDEC and Akron, respectively). These results are encouraging for modeling using decision support systems such as GPFARM in that most varieties tended to respond the same, both between years and locations. Norstar was an anomaly at both locations. Perhaps the percent yield loss response of Norstar, which is a cultivar not typically grown in Colorado, may be partly explained as differentially responding to environmental variables other than water because it is adapted/bred for a different region (e.g., photoperiod and vernalization requirements; heat, cold and drought tolerance; etc.). However, both Arlin and Prowers 99 are commonly grown in Colorado, and this cannot explain their different behavior compared to the other varieties. Further, why did Arlin and Prowers 99 behave differently at only one location? Possibly the unexplained slower rate of Arlin seedling emergence in 2001 at ARDEC caused a greater yield loss in 2001, as the importance of seedling emergence is generally recognized (e.g., McMaster et al., 2002a). In 2001 at Akron, rust was unusually severe and snow occurred on 21 May, resulting in about 5 h below freezing, but nothing unusual in response was observed for Prowers 99.

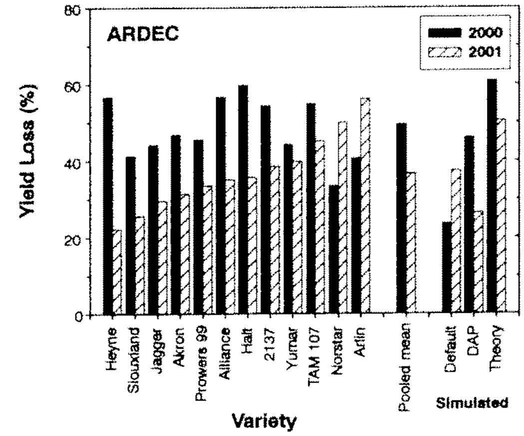

Figure 5. Observed and simulated yield loss for 2000–2001 at ARDEC, Colorado. Simulated values are the three rightmost sets of bars for the three parameters sets evaluated. Percent yield loss = (irrigated – dryland yield)/irrigated yield × 100.

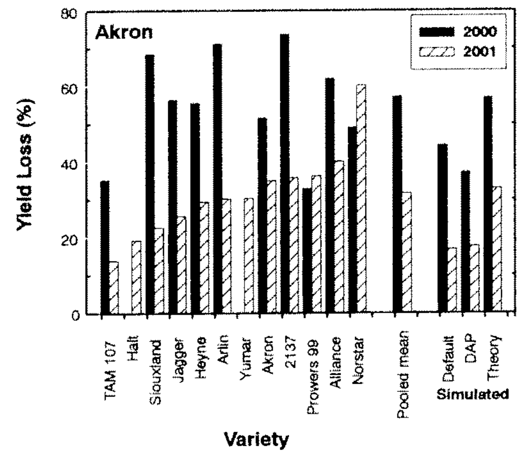

Figure 6. Observed and simulated yield loss for 2000–2001 at Akron, Colorado. Simulated values are the three rightmost sets of bars for the three parameters sets evaluated. Percent yield loss = (irrigated – dryland yield)/irrigated yield × 100. Varieties Halt and Yumar were not grown in 2000.

Regardless of causes, clearly not all varieties responded similarly in yield loss between years or locations, and it is unknown how to explain the observed G × E interaction.

The DAP and Theory parameter sets simulated the pooled mean percent yield loss patterns correctly for both locations (figs. 5 and 6; statistical results not shown). The Default parameter set simulated a very low percent yield loss in 2000 at ARDEC, and therefore the observed pattern between 2000 and 2001 was not correctly simulated. No parameter set was able to correctly simulate both types of varietal yield loss responses observed between years discussed in the preceding paragraph.

The G × E interaction on grain yield can further be seen by examining the change in ranking of percent yield loss of varieties between years (figs. 5 and 6). For instance, at ARDEC, Heyne had among the highest percent yield loss in 2000 (57%), along with Halt (60%), Alliance (57%), TAM 107 (55%), and 2137 (54%), but the lowest percent yield loss

Table 3. Statistical results of simulating TAM 107 grain yield (kg/ha) across locations using the three plant parameter data sets (table 1). Simulation results are presented in figures 9–11.

Statistical Measure[a]	Default	DAP	Theory
RMSE	1031	893	811
SRES	43422	9294	4787
SARES	73676	64745	62068
Paired t–test (P)	<0.0001	0.29	0.55
% of points within 20% of observed yield	40.4	36.2	36.2

[a] RMSE = root mean square error.
SRES = sum of residuals.
SARES = sum of absolute residuals.
P = probability that there is no difference between observed and simulated yield.

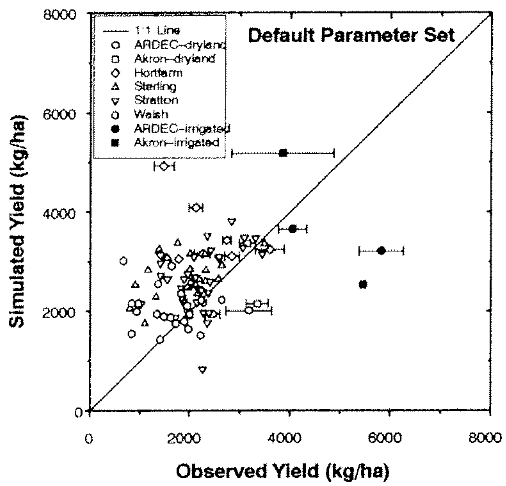

Figure 9. Observed vs. simulated grain yield for cultivar TAM 107 for the ARDEC, Akron, Hortfarm, and DAP (comprised of Sterling, Stratton, and Walsh sites) validation data sets using the Default parameter set. SE bars available only for ARDEC, Akron, and Hortfarm locations. Open symbols are dryland conditions; closed symbols are irrigated conditions.

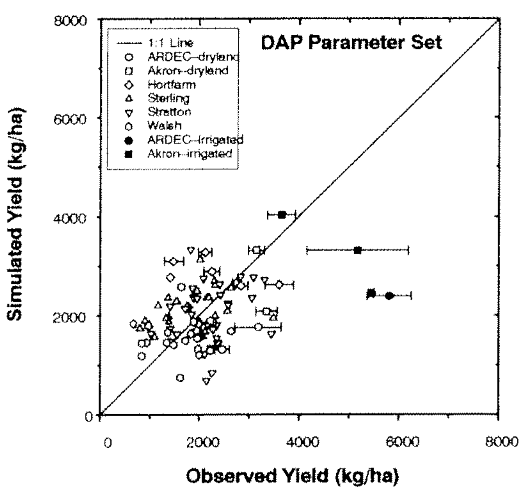

Figure 10. Observed vs. simulated grain yield for cultivar TAM 107 for the ARDEC, Akron, Hortfarm, and DAP (comprised of Sterling, Stratton, and Walsh sites) validation data sets using the DAP parameter set. SE bars available only for ARDEC, Akron, and Hortfarm locations. Open symbols are dryland conditions; closed symbols are irrigated conditions.

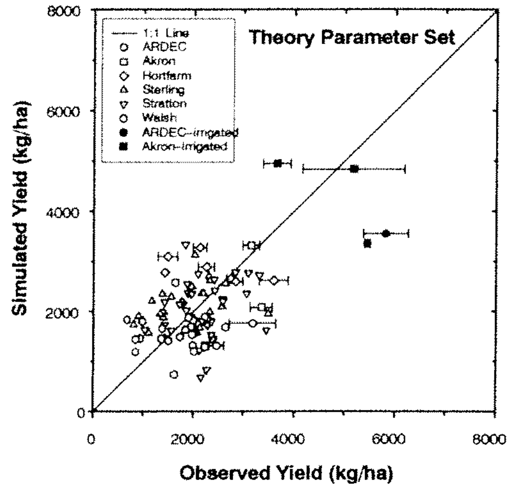

Figure 11. Observed vs. simulated grain yield for cultivar TAM 107 for the ARDEC, Akron, Hortfarm, and DAP (comprised of Sterling, Stratton, and Walsh sites) validation data sets using the Theory parameter set. SE bars available only for ARDEC, Akron, and Hortfarm locations. Open symbols are dryland conditions; closed symbols are irrigated conditions.

served yield, indicated the Default parameter set as being slightly better than the other parameter sets.

TAM 107 was a variety with inconsistent G × E interaction responses to water stress when comparing between locations (figs. 7 and 8). When considering model performance across a range of locations with different management practices, soils, and climates, it appears that a species–based parameter set divided into irrigated and dryland parameter subsets (e.g., Theory parameter set) can simulate TAM 107 grain yield without bias, although whether the scatter around the one–to–one line is acceptable is best left to the individual to decide. Besides the normal causes of simulation error, it is likely that the G × E interaction is significantly contributing to this error. By dividing parameters into two sets for dryland and irrigated conditions, which was easily done *a priori* in our case, a simple alternative was provided for improving plant response relationships with environmental stresses for simulating yield.

Summary and Conclusions

Species–based plant parameter sets cannot reproduce all the complex G × E interactions exhibited by varieties for grain yield. If "unusual" varieties are simulated, the user must be aware that accuracy of results depends greatly on the degree of G × E interaction exhibited by the variety, and certainly significant error is introduced into yield predictions by varietal differences.

Simulating grain yield for one wheat variety (TAM 107), which showed some G × E interaction, across a range of locations and environments was improved by subdividing the parameters into irrigated and dryland values. Undoubtedly, a significant portion of the error was caused by poor mechanistic representation of the G × E interaction in the model. It was not sufficient to merely change parameter values for the potential levels or rates of processes for a variety without also knowing how the level or rate is altered across environmental