EVALUATING RZWQM2-CERES-MAIZE AND WATER PRODUCTION FUNCTIONS FOR PREDICTING IRRIGATED MAIZE YIELD AND BIOMASS IN EASTERN COLORADO

N. Q. Sima, A. A. Andales, R. D. Harmel, L. Ma, T. J. Trout

ABSTRACT. *Complex crop models have been developed to simulate the interactions among biophysical processes and to extend experimental results beyond the local soil and climate conditions. However, in-depth studies on a model's capability to predict crop growth under different conditions are sparse, and the question of whether a crop model outperforms a simple water production function (WPF) has not been answered. The objective of this study was to compare the predictive ability of a complex crop model with simple WPFs for yield and biomass estimation at three sites (Greeley, Fort Collins, and Akron) in eastern Colorado. Specifically, the CERES-Maize crop model in the Root Zone Water Quality Model (RZWQM2), which has been applied extensively in eastern Colorado for simulating maize growth, was compared to crop WPFs based on irrigation and rainfall amounts during growing seasons. Results showed that the predictive ability of CERES-Maize depended on which datasets were used for model parameterization, and that WPFs in general performed as good as or better than CERES-Maize based on a modified F-test after considering experimental uncertainties. The ability of CERES-Maize and the WPF derived from Greeley (2008-2011) to predict maize yield in Greeley (2012-2013), Fort Collins (2006-2010), and Akron (1984-1986) depended on year and site. WPFs outperformed CERES-Maize for Greeley (2012-2013) and Fort Collins (2006-2010) but performed similarly for Akron (1984-1986). This study also identified the need to improve crop model responses to water stress, especially at different growth stages, for cropping systems models to be adequate for estimating the impacts of irrigation management on yield. Ultimately, the choice between the use of a complex crop model and a simpler WPF depends on the purpose of the user and the required accuracy.*

Keywords. *Biomass, CERES-Maize, DSSAT, Grain yield, Irrigation management, RZWQM, Water production function.*

Cropping systems models are developed on the premise that they can be flexible and transferrable across locations and management practices and can be used to simulate the interactions among system components, as compared with site-specific regression equations. Typically, these models require considerable effort to calibrate and have uncertainty in their parameters (O'Grady and O'Hare, 2017; Araujo et al., 2013). Ma et al. (2012a) used the CERES-Maize crop model in the Root Zone Water Quality Model (RZWQM2) to simulate maize yield and biomass responses to irrigation treatments (six irrigation treatments from 2008 to 2011) in the Great Plains and found that parameters derived from one year of experimental data could not predict maize growth in other years unless the model was recalibrated. They suggested using multiple treatments or multiple years of data for model calibration. Later, when Fang et al. (2017) used RZWQM2 at the same experimental site as Ma et al. (2012a) to simulate a maize variety of similar maturity from 2012 to 2013, they had to recalibrate the crop parameters to obtain reasonable maize responses to irrigation based on growth stage. Therefore, it is very common in the literature that a system model was calibrated from site to site, because each site had different soil types or different crop cultivars (Zhang et al., 2018a; Saseendran et al., 2014, 2015a).

On the other hand, regression equations derived from experimental data may be more practical because they involve simpler statistical curve-fitting that may provide predictive accuracy comparable to mechanistic models, especially when uncertainty in the input parameters is considered (Araujo et al., 2013). Trout and DeJonge (2017) found that regression equations for water production functions (WPFs) might be adequate for practical purposes and that the results could be scaled among irrigation treatments and across years when maize yield was a function of actual crop evapotranspiration (ET). Nielsen et al. (2011) used WPFs for several

Submitted for review in August 2018 as manuscript number NRES 13045; approved for publication by the Natural Resources & Environmental Systems Community of ASABE in December 2018.

Mention of trade names is solely for providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer.

The authors are **Nathan Q. Sima,** Student Intern, and **Allan A. Andales,** Associate Professor, Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado; **R. Daren Harmel,** Director, USDA-ARS Center for Agricultural Resources Research, Fort Collins, Colorado; **Liwang Ma,** Soil Scientist, USDA-ARS Rangeland Resources and Systems Research Unit, Fort Collins, Colorado; **Thomas J. Trout,** Agricultural Engineer (Retired), USDA-ARS Water Management and Systems Research Unit, Fort Collins, Colorado. **Corresponding author:** Allan A. Andales, 1170 Campus Delivery, Fort Collins, CO 80523-1170; phone: 970-491-6516; e-mail: allan.andales@colostate.edu.

crops along with estimated growing season precipitation and estimated growing season soil water extraction to estimate yields to determine which crops to plant in dryland crop rotations in northeast Colorado. Additionally, Nielsen et al. (2012) presented maximum, minimum, and mean estimated canola yields from a previously published production function and from a cropping systems model that produced mostly similar results. They found, however, that the maximum estimated canola yield was greater with the production function because it did not account for yield limitations that may have resulted from high-temperature stress.

However, WPFs can vary greatly from year to year and from location to location (Nielsen et al., 2005; Nielsen and Vigil, 2017a, 2017b). Other factors may need to be included in the regression equations to improve model predictive ability (Nielsen et al., 2017b). In such cases, system models can be used to derive regression-type equations based on simulation results that incorporate major agronomic factors and their interactions. For example, Saseendran et al. (2015b) used simulation results from 1992-2011 using the calibrated CERES-Maize crop model in RZWQM2 to develop maize WPFs at three sites (Greeley, Akron, and Rocky Ford) in Colorado and one site in China and found that the simulated results could be scaled across sites and soil types when relative yield and relative ET were used. However, it has yet to be determined whether system models or empirical regression-type WPFs have better predictive ability and transferability across sites and irrigation conditions. To explore this issue, the objective of this study was to compare the CERES-Maize crop model in RZWQM2 with simple regression-type crop WPFs for their ability to predict irrigated maize biomass and grain yield at three sites (Greeley, Fort Collins, and Akron) in eastern Colorado.

Materials and Methods

Uniform Water Stress Experiments at Greeley

The uniform water stress experimental data were obtained from a study conducted from 2008 to 2011 near Greeley, Colorado (40.45° N, 104.64° W). The soil is a sandy loam and is uniform throughout the 200 cm soil profile. Maize (Dekalb 52-59, 102-day maturity class) was planted at an average rate of 81,000 seeds ha^{-1} with 0.76 m row spacing in early May annually from 2008 to 2011. The field was divided into four replicated blocks, and each block was divided into twelve 9 m × 44 m plots. Weather data were recorded on-site with a standard Colorado Agricultural Meteorological Network (http://ccc.atmos.colostate.edu/~coagmet/) evapotranspiration weather station and were used to calculate ASCE Standardized Penman-Monteith alfalfa reference ET (ET$_r$) (Allen et al., 2005). Irrigations were scheduled based on crop ET (ET$_c$), which was estimated from ET$_r$ and FAO-56 dual crop coefficients under well irrigated conditions (Allen et al., 1998; Comas et al., 2019). Six irrigation treatments (micro-irrigation with surface drip tubing adjacent to each row) were designed to meet a specified percentage of ET$_c$ (Allen et al., 1998, 2005) during the growing seasons: 100% (T1), 85% (T2), 75% (T3), 70% (T4), 55% (T5), and 40% (T6) of ET$_c$. The amount of water for each treatment was estimated at a 3 to 6 day interval based on ET$_c$, rainfall, and soil water deficit (Trout and Bausch, 2017). The T1 treatment was irrigated such that water availability (irrigation plus precipitation plus stored soil water) was adequate to meet ET$_c$. The remaining treatments were irrigated to meet a certain percentage of water demand in T1. All treatments received full irrigation as T1 up to the seven-leaf vegetative stage (V7). Water stress was partially alleviated again during the reproductive stage (VT to R2) to ensure adequate pollination and seed initiation (Trout and Bausch, 2017; Trout and DeJonge, 2017). Growth stages were recorded twice a week with standard notation of stages for maize (Abendroth et al., 2011).

Soil water content was measured two or three times each week in 30 cm depth increments between 30 and 150 cm depth, and at 200 cm depth, with a neutron soil moisture meter (CPN-503 Hydroprobe, InstroTek, San Francisco, Cal.). Soil water content in the top 15 cm was measured in the row near the neutron moisture meter access tube with a portable time domain reflectometer (TDR, Minitrase, Soilmoisture Equipment Corp., Santa Barbara, Cal.). Aboveground biomass was measured four times per season by collecting ten plants from the middle rows. Final yield was also measured by hand-harvesting the ears from the middle 15 m of the center four rows of each plot (Trout and Bausch, 2017; Comas et al., 2019). The experimental dataset and detailed methodology are available at the USDA National Agricultural Library Ag Data Commons (http://dx.doi.org/10.15482/USDA.ADC/1254006) (Trout and Bausch, 2017; Trout and DeJonge, 2017, 2018).

Growth-Based Water Stress Experiments at Greeley

Water stress experimental data based on growth stage were obtained at the same site from 2012 to 2013 with the same experimental methodology as used in 2008-2011, except for irrigation scheduling. Maize (Dekalb DKC52-04RIB, also a 102-day maturity class as Dekalb 52-59) was planted at an average rate of 84,000 seeds ha^{-1} in 2012 and 85,500 seeds ha^{-1} in 2013 at 0.76 m spacing (Trout and DeJonge, 2018; Comas et al., 2019). Different from the 2008-2011 experiments, irrigation was scheduled based on growth stage. Twelve irrigation treatments were randomly assigned to the plots within each block (randomized complete block). The irrigation treatments were designed to meet various target percentages of ET$_c$ during the vegetative (V7 to VT) and late reproductive (R4 to R6) growth stages and full ET$_c$ during the early reproductive stage (R1 to R3). The target percentages of full ET$_c$ were 40% to 100% at the V7 to VT and R4 to R6 stages for these irrigation treatments (e.g., 65/40 indicates 65% of full ET$_c$ during V7-VT and 40% of full ET$_c$ during R4-R6). Twelve treatments were created and denoted as (targeted ET$_c$ at vegetative stage)/(targeted ET$_c$ at reproductive stage) (100/100, 100/50, 80/80, 80/65, 80/50, 80/40, 65/80, 65/65, 65/50, 65/40, 50/50, and 40/40). Other details are available from Fang et al. (2017) and Comas et al. (2019), and the experimental dataset and detailed methodology are available at the USDA National Agricultural Library Ag Data Commons (http://dx.doi.org/10.15482/USDA.ADC/1439968).

Model Calibration and WPF Generation at Greeley

The CERES-Maize model in RZWQM2 was selected because it provided similar results as in DSSAT (Decision Support System for Agrotechnology Transfer) (Ma et al., 2006) and because its parameters could be optimized objectively with the Parameter ESTimation (PEST) tool (Ma et al., 2012a; Sima et al., 2018). The 2008-2011 and 2012-2013 data were simulated and published in previous publications using CERES-Maize in RZWQM2 (Sima et al., 2018; Ma et al., 2012a; Fang et al., 2017; Saseendran et al., 2014, 2015a, 2015b). Details on model parameterization are available from Sima et al. (2018) for 2008-2011 data and from Fang et al. (2017) for 2012-2013 data. Specifically, PEST was used to optimize maize cultivar parameters using data from each year individually (i.e., 2008, 2009, 2010, 2011) or from all four years (2008-2011). These sets of cultivar parameters were then used to predict maize yield and biomass in 2008-2011 and in 2012-2013. All model runs started on January 1 each year. Similar to model calibration, crop WPFs were developed from the data for each year (i.e., 2008, 2009, 2010, 2011) or for all years combined (2008-2011) by linear regression with total water input (rainfall + irrigation) during each growing season (May to October) as the independent variable because irrigation and rainfall accounted for more than 90% of crop ET (Trout and Bausch, 2017). In addition, it would be easier for farmers to obtain rainfall and irrigation data than to estimate ET_c. These WPFs were then used to predict maize yield and biomass in 2008-2011 (uniform irrigation) and in 2012-2013 (growth-stage based irrigation).

Prediction of Maize Yield at Fort Collins and Akron Sites

The plant cultivar parameters and WPF derived from the 2008-2011 data were further evaluated for their predictions of maize yield at Fort Collins (40.65° N, 105.00° W) and Akron (40.15° N, 103.14° W), Colorado, regardless of differences in maize cultivars. The data for Fort Collins were collected from 2006 to 2010 at Colorado State University's Agricultural Research Development and Education Center (ARDEC) and were published by DeJonge et al. (2011, 2012) and Zhang et al. (2018b). The Akron data were from 1984 to 1986 at the USDA-ARS Central Great Plains Research Station with four irrigation treatments and were published by Ma et al. (2003) and Saseendran et al. (2014, 2015a, 2015b). The Fort Collins data had one fully irrigated and one limited irrigation treatment, and the limited irrigation was only applied before the V12 stage (DeJonge et al., 2011, 2012). After the V12 stage, there was no water stress in both treatments. On the other hand, irrigation in the Akron study was initiated prior to tasseling (VT stage) in each year (Ma et al., 2003). Therefore, the low-irrigation treatments could have been under water stress during the tasseling and grain filling stages in the Akron study. In addition, none of the treatments were under full irrigation in Akron. On-site weather data were collected at both sites, and soil hydrological data were from DeJonge et al. (2011) for Fort Collins and from Ma et al. (2003) for Akron. Summary information of the experiments at the three sites is available from Zhang et al. (2018b) and is listed in table 1. Model runs started on January 1 each year. Because no nitrogen stress was observed in the experiments, the model was run without N stress. For WPF prediction, no adjustment was made for population differences among the sites.

Statistics for Model Comparison

Objective model performance comparison, considering experimental errors, has been recently recommended using a modified version of the F-test (Sima et al., 2018). The modified F-test is defined as (Sima et al., 2018; Ma et al., 2012b):

$$F = \frac{\text{MSLOFIT}}{\text{MSE}} \tag{1}$$

where

$$\text{MSLOFIT} = \frac{\sum_{i=1}^{L} N_i \left(P_i - O_i\right)^2}{\sum_{i=1}^{L} N_i} \tag{2a}$$

Table 1. Crop management practices of irrigated maize at three sites (Greeley, Fort Collins, and Akron) in eastern Colorado (Adapted from Zhang et al., 2018b).

Site	Year (Treatments)	Cultivar	Plant Population (plants ha^{-1})	Planting Date	Harvest Date	Precipitation (May-Oct.) (mm)	Irrigation Amount for Each Treatment (mm)
Greeley	2008 (6)	Dekalb 52-59	81,000	May 12	Nov. 6	250	438, 338, 282, 272, 181, 137
	2009 (6)	Dekalb 52-59	81,000	May 11	Nov. 12	225	418, 348, 300, 250, 168, 109
	2010 (6)	Dekalb 52-59	81,000	May 11	Oct. 19	211	365, 290, 247, 220, 159, 112
	2011 (6)	Dekalb 52-59	81,000	May 3	Oct. 25	201	485, 388, 329, 306, 221, 157
	2012 (12)	Dekalb 52-04	84,000	April 30	Oct. 19	145	698, 523, 583, 527, 479, 433, 578, 478, 458, 459, 427, 365
	2013 (12)	Dekalb 52-04	85,500	May 15	Nov. 4	250	509, 415, 425, 404, 373, 345, 402, 383, 353, 326, 325, 295
Fort Collins	2006 (2)	Garst 8827	79,100/59,300[a]	May 10	Nov. 4	868	500, 259
	2007 (2)	Garst 8827	79,800/59,300[a]	May 10	Nov. 14	201	362, 210
	2008 (2)	Pioneer 38P	79,100	April 30	Nov. 19	245	406, 203
	2009 (2)	Pioneer P9512XR	79,100	May 13	Nov. 13	201	292, 191
	2010 (2)	Producers Hybrids 5004VT3	79,100	May 4	Oct. 16	245	400, 210
Akron	1984 (3)	Pioneer 3732	72,400	May 14	Oct. 1	342	106, 68, 23
	1985 (4)	Pioneer 3732	76,100	May 3	Sept. 17	340	188, 151, 98, 72
	1986 (4)	Pioneer 3732	76,100	May 1	Oct. 15	231	299, 258, 203, 146

[a] Values are for full irrigation / limited irrigation.

$$\text{MSE} = \frac{\sum_{i=1}^{L}\sum_{j=1}^{N_i}\left(O_{ij} - O_i\right)^2}{\sum_{i=1}^{L}\left(N_i - 1\right)} \tag{2b}$$

with degrees of freedom of:

$$\nu_1 = \sum_{i=1}^{L} N_i \text{ for the numerator}$$

$$\nu_2 = \sum_{i=1}^{L}\left(N_i - 1\right) \text{ for the denominator}$$

L = number of measurement groups, which may represent different treatments or different sampling dates

N_i = number of measured replicates for ith measurement group

P_i = predicted value for ith measurement group

O_{ij} = jth observation (replicate) for ith measurement group

O_i = mean of ith experimental or measurement group:

$$O_i = \frac{1}{N_i}\sum_{j=1}^{N_i} O_{ij}$$

The F-test calculates a p-value from the F probability distribution that is used to test the null hypothesis that the experimental values are the same as the simulated values (Ma et al., 2012b). Using confidence level $\alpha = 0.05$, if $p > 0.05$, then simulated values are not significantly different from observed values.

In addition to the modified F-test, three other statistics were used as comparison: RRMSE (relative root mean squared error), RSpR (RMSE to average experimental error ratio), and r^2 (coefficient of determination) (Ma et al., 2012b):

$$\text{RRMSE} = \frac{\sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(P_i - O_i\right)^2}}{\frac{1}{L}\sum_{i=1}^{L} O_i} \tag{3}$$

$$\text{RSpR} = \frac{\sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(P_i - O_i\right)^2}}{S_p} \tag{4}$$

where S_p is the pooled standard derivation or experimental error, as:

$$S_p^2 = \frac{\sum_{i=1}^{L}\sum_{j=1}^{N_i}\left(O_{ij} - O_i\right)^2}{\sum_{i=1}^{L}\left(N_i - 1\right)} \tag{5}$$

$$r^2 = \frac{\left[\sum_{i=1}^{L}\left(O_i - O_{avg}\right)\left(P_i - P_{avg}\right)\right]^2}{\sum_{i=1}^{L}\left(O_i - O_{avg}\right)^2 \sum_{i=1}^{L}\left(P_i - P_{avg}\right)^2} \tag{6}$$

where $O_{avg} = \frac{1}{L}\sum_{i=1}^{L} O_i$ and $P_{avg} = \frac{1}{L}\sum_{i=1}^{L} P_i$

Ma et al. (2011, 2012b) indicated satisfactory model prediction when $r^2 > 0.80$, $\text{RSpR} < 1.0$, and $\text{RRMSE} < 0.10$ for yield and biomass simulations.

Results and Discussion

Calibrated crop cultivar parameters of CERES-Maize varied greatly among the sub-datasets used for optimization (table 2). Among the four years (2008, 2009, 2010, and 2011), P1 (thermal time from seedling emergence to end of juvenile phase) ranged from 244 to 308, P2 (day length sensitivity coefficient) ranged from 0.22 to 0.52, P5 (degree days from silking to physiological maturity) ranged from 592 to 1000, G2 (potential kernel number per plant) ranged from 460 to 990, G3 (potential kernel growth rate) ranged from 6.1 to 16.0, and PHINT (thermal time required for a leaf tip to emerge) ranged from 38.0 to 42.5. There was a strong influence of the experimental dataset used for calibration on the calibrated cultivar parameters (Sima et al., 2018). For regressed WPFs, the slope for seasonal biomass varied from 29.5 to 38.4 kg ha^{-1} per mm water, and the intercepts ranged widely from -3421 to 2307 kg ha^{-1}. Similarly, the slope for grain yield varied from 14.2 to 24.9 kg ha^{-1} per mm water, and the intercepts ranged widely from -4685 to 2120 kg ha^{-1}. There was a strong correlation between slope and intercept: more negative intercepts corresponded to higher slopes (table 2).

When experimental uncertainty was considered, seasonal biomass from 2008-2011 was predicted well by CERES-Maize ($p > 0.05$, suggesting no significant difference between simulated and measured results) when maize cultivar parameters were obtained from 2011 data only or from all 2008-2011 data (table 3). On the contrary, simulated biomass from CERES-Maize with cultivar parameters calibrated from 2009 and 2010 data were significantly different from measured biomass ($p < 0.05$). CERES-Maize with cultivar parameters from 2008 data severely overpredicted biomass in 2011, and the model with parameters from 2009 data significantly underpredicted biomass in all other years (fig. 1). On the other hand, all WPFs predicted 2008-2011 seasonal biomass well except for that derived from 2011 data, which often underpredicted biomass in other years (table 3). In general, the biomass predictions from CERES-Maize varied more from year to year than those from WPFs (fig. 1).

Similar results were obtained for simulated maize grain yield (table 4), with cultivar parameters from 2011 data only or from all 2008-2011 data providing satisfactory yield prediction when experimental uncertainty was considered. On

Table 2. Optimized maize cultivar parameters (CERES-Maize) and water production functions (WPFs) derived from data of all six treatments in 2008, 2009, 2010, 2011, or 2008-2011 at the Greeley, Colorado, site (Sima et al., 2018).

Optimized Maize Cultivar Parameters in CERES-Maize or WPFs		Parameters Fitted from All Treatments in Each Year or Multiple Years				
		2008	2009	2010	2011	2008-2011
P1	Degree days (base temperature of 8°C) from seedling emergence to end of juvenile phase (thermal degree days).	307.8	236.8	244.4	275.0	263.9
P2	Day length sensitivity coefficient (extent, in days, by which development is delayed for each hour increase in photoperiod above the longest photoperiod (12.5 h) at which development proceeds at the maximum rate.	0.38	0.26	0.27	0.52	0.22
P5	Degree days (base temperature of 8°C) from silking to physiological maturity (thermal degree days).	1000.0	591.8	568.6	629.7	725.0
G2	Potential kernel number per plant.	459.9	971.8	470.5	989.8	554.3
G3	Potential kernel growth rate (mg kernel^{-1} d^{-1}).	13.31	6.08	16.00	16.00	10.42
PHINT	Degree days required for a leaf tip to emerge (thermal degree days).	39.5	39.6	55.0	38.0	42.5
	r^2 for yield simulation	0.74	0.56	0.87	0.94	0.95
	r^2 for biomass simulation	0.83	0.84	0.88	0.89	0.91
Regression equation (WPF) for biomass: Y (kg ha^{-1}) = aX (mm rainfall + irrigation) + b	Intercept (b)	2307.7	1803.3	-2196.7	-3421.6	-369.5
	Slope (a)	29.5	31.9	37.7	38.4	34.3
	r^2	0.88	0.90	0.95	0.97	0.89
Regression equation (WPF) for yield: Y (kg ha^{-1}) = cX (mm rainfall + irrigation) + d	Intercept (d)	2119.2	66.4	-956.3	-4685.5	-827.3
	Slope (c)	14.2	16.9	19.6	24.9	18.8
	r^2	0.85	0.91	0.90	0.94	0.84

Table 3. F-test, RSpR, RRMSE, and r^2 for predicting seasonal maize biomass in 2008-2011 and 2012-2013, using crop parameters (CERES-Maize) or water production functions (WPFs) derived from datasets in 2008, 2009, 2010, 2011, and 2008-2011 at the Greeley, Colorado, site.[a]

Model	Data Used for Prediction	Data Used for Calibration	RSpR	RRMSE	r^2	F-Value	p-Value
CERES-Maize	2008-2011	2008	1.416	0.119	0.828	2.005	0.001
		2009	2.061	0.173	0.841	4.250	<0.001
		2010	1.644	0.138	0.881	2.704	<0.001
		2011	1.114	0.094	0.889	1.242	0.167
		2008-2011[b]	0.828	0.070	0.911	0.704	0.950
	2012-2013	2008-2011	0.952	0.114	0.595	0.906	0.677
WPF	2008-2011	2008	0.995	0.084	0.885	0.990	0.521
		2009	0.981	0.082	0.885	0.963	0.572
		2010	1.056	0.089	0.885	1.115	0.316
		2011	1.451	0.122	0.885	2.105	<0.001
		2008-2011[b]	0.992	0.083	0.885	0.984	0.532
	2012-2013	2008-2011	1.152	0.139	0.661	1.327	0.104

[a] RMSE = root mean squared error, RSpR = RMSE to average experimental error ratio, RRMSE = relative RMSE; $p \geq 0.05$ indicates no significant difference between simulated and observed values. All WPF predictions have the same r^2 value because they are linear functions of the same (rainfall + irrigation) water inputs from 2008-2011 (table 1), and thus the r^2 values are independent of WPF slopes and intercepts.

[b] Indicates that the same data were used in both model calibration and prediction.

the contrary, WPF derived from 2011 data did not provide satisfactory yield prediction ($p < 0.05$, table 4). CERES-Maize with cultivar parameters optimized from 2008 data severely underpredicted yield in 2009 and overpredicted yield in 2011 (fig. 2). Cultivar parameters from 2009 data underpredicted yield in all other years (fig. 2). The 2009 crop sustained hail damage during two different growth phases: seedling phase (May 27 and June 1) and tasseling (July 29 and 31). Nevertheless, the crop recovered from the hail damage, and final yields were comparable to other years. However, the combination of optimized parameters derived from the 2009 data set (table 2) was biased toward lower yield predictions in other years (fig. 2). In particular, the combination of shorter thermal degree days from silking to physiological maturity (P5 = 591.8°C d) and low potential kernel growth rate (G3 = 6.08 mg kernel^{-1} d^{-1}) may have led to less accumulation of grain mass in other years. The WPF from the 2011 data considerably underpredicted maize yield in all other years (fig. 2). Again, predictions from CERES-Maize varied more among years than those from WPFs. It is also interesting to note that the predictions were more scattered at higher ET_c treatments for CERES-Maize and at lower ET_c treatments for WPFs, especially for parameters derived from 2011 data.

With cultivar parameters and WPF derived from all the 2008-2011 data, we predicted maize growth in 2012-2013 when irrigation was scheduled based on growth stages. CERES-Maize with these cultivar parameters provided satisfactory prediction for 2012-2013 biomass ($p < 0.05$, table 3, fig. 4). However, the model missed the biomass decrease from the 100/100 treatment to the 100/50 treatment and the increase from 80/40 to 65/80 in 2012 (fig. 3), whereas the WPF simulated these trends well. In 2013, CERES-Maize also failed to predict the high biomass under the 65/80 and 65/65 treatments (fig. 3). Based on the F-test, grain yield predictions for 2012-2013 by the WPF were satisfactory, but not the yield predictions by CERES-Maize (table 4, fig. 3). CERES-Maize could not simulate the yield increase for the 80/80 and 65/80 treatments in both 2012 and 2013 (fig. 3). Between the two years, CERES-Maize's yield predictions were better for 2013 ($p = 0.453$) than for 2012 ($p = 0.00214$), especially for the low-irrigation treatments. On the other hand, WPF overpredicted grain yield for the

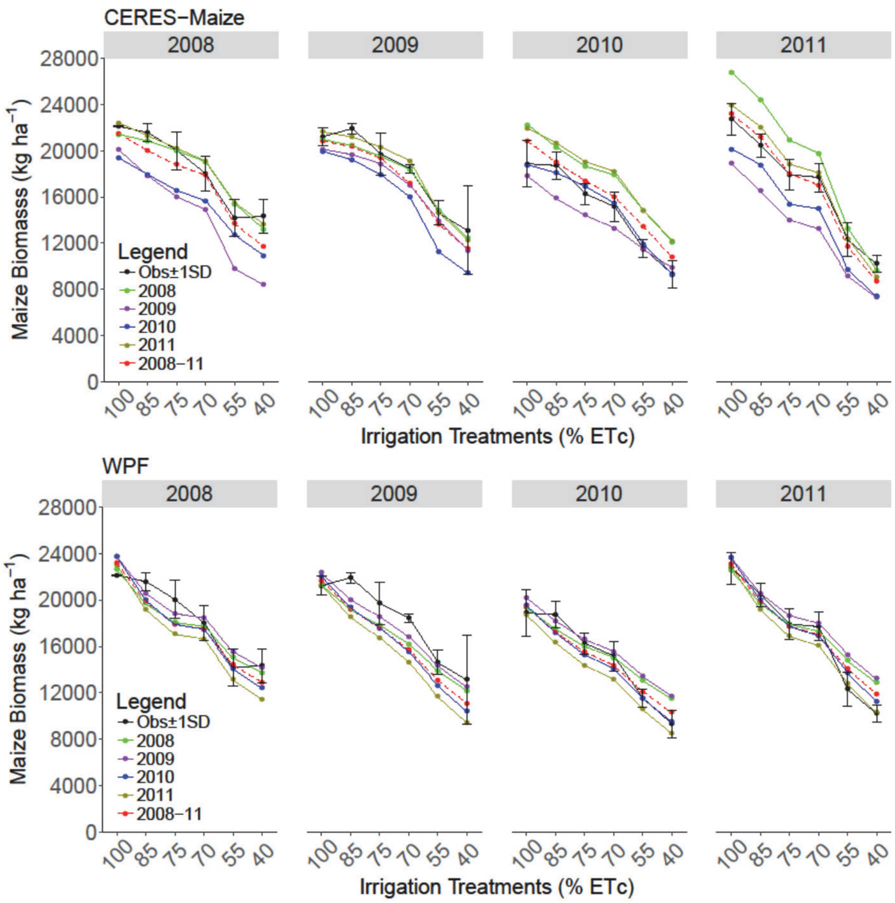

Figure 1. Measured and predicted maize biomass from 2008-2011 using calibration from 2008, 2009, 2010, 2011, and 2008-2011 at the Greeley, Colorado, site. Treatments indicate percentages of crop evapotranspiration (ET_c) supplied by irrigation, rain, and stored soil water. Obs±1SD is measured value ± one standard derivation from the mean. Each color line represents prediction with cultivar parameters (CERES-Maize) or water production functions (WPFs) from an individual year (2008, 2009, 2010, 2011) or all four years (2008-2011).

Table 4. F-test, RSpR, RRMSE, and r^2 for predicting maize yield over various periods at three eastern Colorado sites (Greeley, Fort Collins, and Akron) using crop parameters (CERES-Maize) or water production functions (WPFs) derived from data sets in 2008, 2009, 2010, 2011, or 2008-2011 at the Greeley, Colorado, site.[a]

Model	Data Used for Prediction	Data Used for Calibration	RSpR	RRMSE	r^2	F-Value	p-Value
CERES-Maize	2008-2011 (Greeley)	2008	2.055	0.173	0.742	4.224	<0.001
		2009	3.913	0.330	0.564	15.312	<0.001
		2010	1.154	0.128	0.872	2.292	<0.001
		2011	1.152	0.097	0.937	1.327	0.104
		2008-2011[b]	0.766	0.065	0.954	0.586	0.990
	2012-2013 (Greeley)	2008-2011	1.454	0.204	0.424	2.117	<0.001
	2006-2010 (Fort Collins)	2008-2011	1.357	0.261	0.401	1.841	0.042
	1984-1986 (Akron)	2008-2011	3.413	0.375	0.176	11.647	<0.001
WPF	2008-2011 (Greeley)	2008	1.599	0.135	0.851	2.556	<0.001
		2009	1.199	0.101	0.851	1.438	0.053
		2010	1.206	0.102	0.851	1.455	0.048
		2011	1.896	0.160	0.851	3.595	<0.001
		2008-2011[b]	1.157	0.098	0.851	1.340	0.097
	2012-2013 (Greeley)	2008-2011	0.942	0.132	0.525	0.887	0.710
	2006-2010 (Fort Collins)	2008-2011	1.107	0.196	0.704	1.034	0.467
	1984-1986 (Akron)	2008-2011	1.940	0.213	0.724	3.761	<0.001

[a] RMSE = root mean squared error, RSpR = RMSE to average experimental error ratio, RRMSE = relative RMSE; $p \geq 0.05$ indicates no significant difference between simulated and observed values. All WPF predictions have the same r^2 value because they are linear functions of the same (rainfall + irrigation) water inputs from 2008-2011 (table 1), and thus the r^2 values are independent of WPF slopes and intercepts.

[b] Indicates the same data were used for both model calibration and prediction.

100/100 treatment in both years, which could be due to water loss to deep drainage (Trout and DeJonge, 2017) that was unaccounted for in the WPF. Therefore, if RZWQM2-CERES-Maize, calibrated with the 2008-2011 data, was used for decision-making to reduce irrigation water use during the maturation phase, it would simulate higher biomass and yield and falsely predict economic gain. Overall, the WPFs performed better for predicting biomass and grain

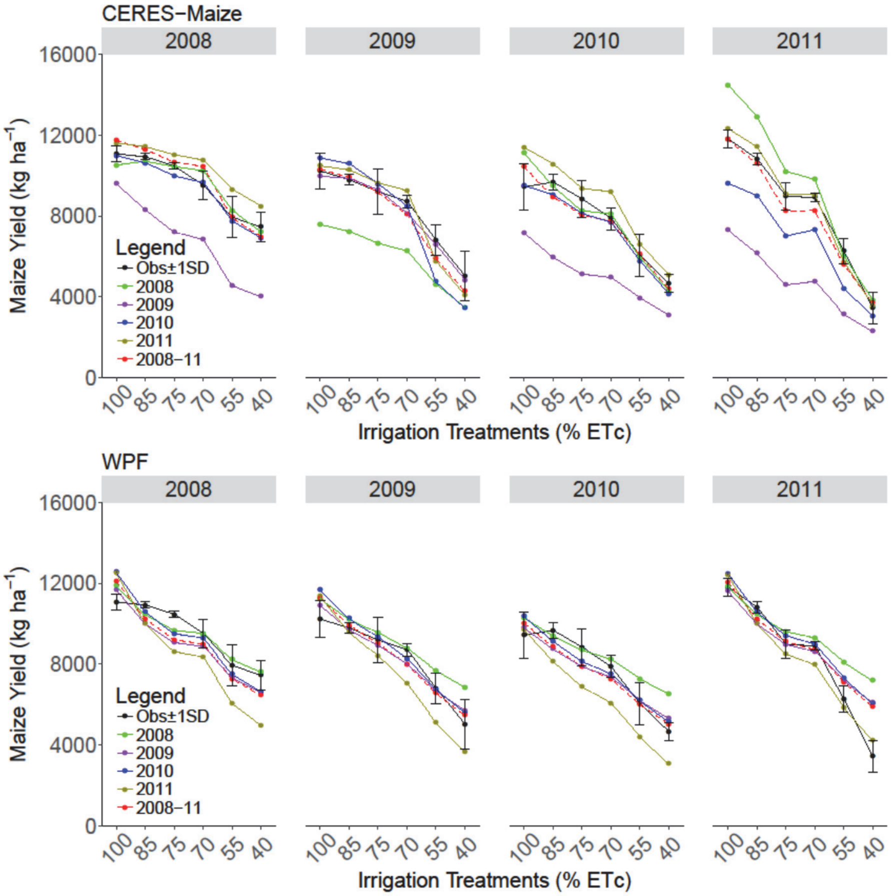

Figure 2. Measured and predicted maize yield from 2008-2011 using calibration from 2008, 2009, 2010, 2011, and 2008-2011 at the Greeley, Colorado, site. Treatments indicate percentages of crop evapotranspiration (ET_c) supplied by irrigation, rain, and stored soil water. Obs±1SD is measured value ± one standard derivation from the mean. Each color line represents prediction with cultivar parameters (CERES-Maize) or water production functions (WPFs) from an individual year (2008, 2009, 2010, 2011) or all four years (2008-2011).

yield in 2012-2013 based on the F-test, RSpR, and RRMSE (tables 3 and 4). At a confidence level of $\alpha = 0.05$ and experimental coefficient of variation (CV) of 0.075 observed in this study, we interpreted RRMSE ≤ 0.1 and RSpR ≤ 1.2 as "very good" simulation results with $p \geq 0.05$ (table 3). However, the critical RRMSE and RSpR values for accepting a model's performance could be higher if experimental errors (CVs) are larger (e.g., CV = 0.10 or 0.15).

Figure 4 shows predicted maize yield at the Fort Collins and Akron sites in Colorado. Statistically, WPF ($p = 0.467$) predicted yield better than CERES-Maize ($p = 0.042$) for the Fort Collins site. For the Akron site, CERES-Maize simulated yield better than WPF in 1984, but the opposite was true for 1985 and 1986 (fig. 4). Across the three years at Akron, neither CERES-Maize nor WPF provided satisfactory results based on p-values ($p < 0.001$) (table 4). This may have been due to significant differences in maize cultivars between the calibration data (2008-2011) at Greeley and the test data (1984-1986) at Akron and demonstrates the need for caution when using crop model parameters or WPFs to estimate yields across locations. However, when 1984 data were excluded, the WPF predictions for 1985 and 1986 were acceptable ($p = 0.10$).

When yield data and predictions were lumped across all three sites (2012-2013 at Greeley, 2006-2010 at Fort Collins, and 1984-1986 at Akron), the WPF provided satisfactory results ($p = 0.18$, $r^2 = 0.55$, RRMSE = 0.16, and RSpR = 1.08), whereas CERES-Maize did not ($p < 0.001$, $r^2 = 0.49$, RRMSE = 0.25, RSpR = 1.66; statistics in addition to those shown in table 4). However, both CERES-Maize and WPF had similar r^2 values when the regression intercept was set to zero (fig. 5). With non-zero intercepts, the regression equations between predicted and measured maize yields indicated a bias toward overpredicting low yields. The WPF had a greater tendency to overpredict low yields compared to CERES-Maize due to higher intercept (fig. 5; 2976 kg ha^{-1} vs. 1765 kg ha^{-1}).

It should be noted that simulation results from CERES-Maize depend on simulated soil hydrology. In general, RZWQM2 simulated higher evaporation than estimated from the soil water balance (Trout and DeJonge, 2017) and soil moisture monitoring (Fang et al. 2010). The average annual soil evaporation was 230 mm for Greeley, 250 mm for Fort Collins (ARDEC), and 230 mm for Akron. Negligible deep drainage was simulated for Greeley and Akron, but 54 mm was simulated for the Fort Collins site. Simulated surface runoff was 3 to 19 mm for Greeley, 25 mm for Fort Collins, and 7 to 20 mm for Akron. Crop transpiration ranged from 270 to 620 mm at Greeley, from 380 to 550 mm at Fort Collins, and from 270 to 430 mm at Akron. Because soil evaporation was a major water loss in the RZWQM2

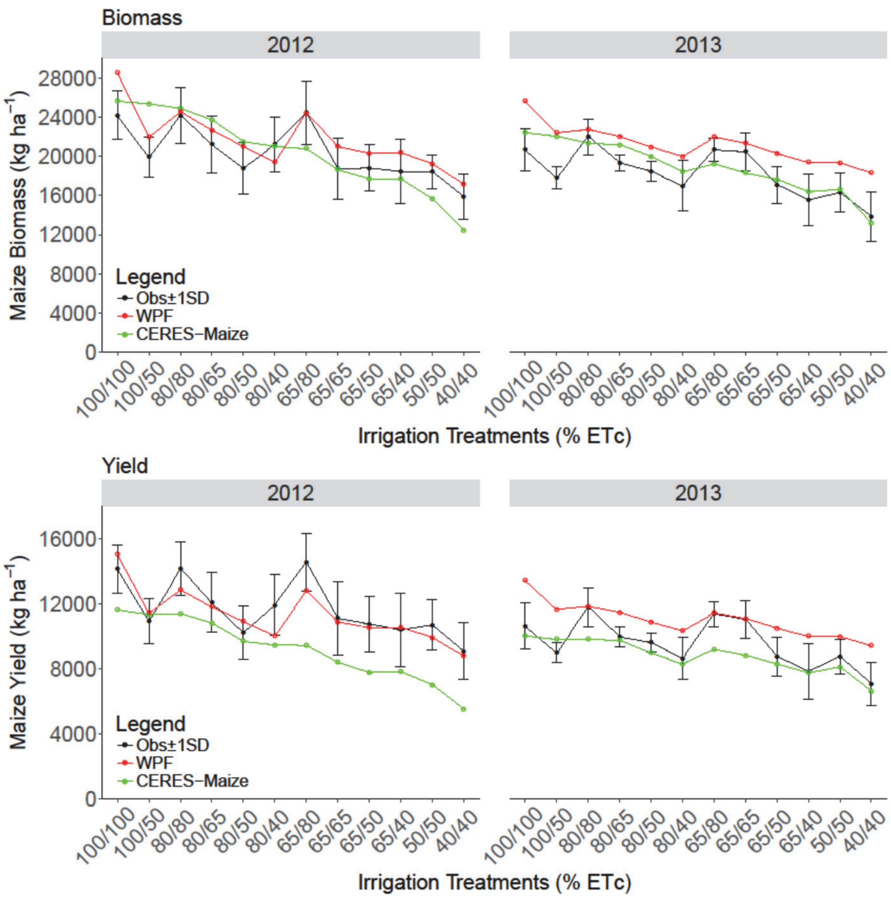

Figure 3. Measured and CERES-Maize predicted harvested maize biomass and yield for the 2012-2013 data using calibration from 2008-2011 at the Greeley, Colorado, site. Treatments denote target percentages of crop evapotranspiration (ET_c) during vegetative (V7 to VT) and late reproductive (R4 to R6) growth stages. Obs±1SD is measured value ± one standard derivation from the mean.

simulations, correct simulation of this component should improve crop simulation. Nonetheless, simulated transpiration by RZWQM2 was close to field estimates for the Greeley site (Ma et al., 2012a).

In addition to soil hydrology, other factors affecting the performance of cropping systems models include (1) the empirical nature of simulating system components and their interactions, (2) the lack of independence among model parameters (i.e., correlations between parameters), and (3) the tendency of fitted model parameters to compensate for model deficiencies (i.e., insufficient plant responses to water stress). Lack of data for calibrating individual processes typically leads to over-parameterization and poor model performance across sites and conditions (Getz et al., 2018). Model parameters are usually obtained through an optimization algorithm (e.g., PEST) or by trial-and-error using a limited data set. Thus, calibrated parameters for complex crop models (e.g., CERES-Maize) may depend on the datasets used and may be site-specific as well (Ma et al., 2011, 2012a, 2012b; Ahuja and Ma, 2011; Seidel et al., 2018).

On the other hand, WPFs are easy to use and easily understood by producers. From the slope of the regression equation, farmers can estimate the grain water use efficiency (kg ha^{-1} per mm H_2O) of their water management practices, and the intercept of the regression equation can indicate the minimum water input (mm rainfall + irrigation) needed to obtain a measurable yield. There is no calibration for WPFs. Trout and DeJonge (2017) suggested that this type of equation can be satisfactorily transferred to other soil and climate conditions if the WPF is normalized relative to maximum crop ET and maximum yield and is adjusted for runoff and deep percolation, as proposed by Doorenbos and Kassam (1986). As shown in a recent study by Zhang et al. (2018b), a simple FAO-style soil water stress factor could be adequate for simulating maize production at all three locations. However, it should be noted that, in this study, the WPFs among years at the Greeley site were similar, and a single WPF may be used for all years (Trout and DeJonge, 2017), which may have contributed to the good predictions using WPF. A key feature of the experimental design was that water stress during reproductive stages (VT to R2) was reduced to minimize its effects on pollination and seed initiation. As such, experimental water stress effects were on vegetative growth and grain filling only. This was why the WPF derived from the Greeley site was able to predict yield well at the Fort Collins site, where water stress was not applied during the reproductive stages. Under dryland conditions, water stress could affect all crop growth phases. This was why the WPF did not simulate yield well at the Akron site in 1984, as maize was subjected to water stress during pollination and seed initiation stages. As a result, the actual WPF varied greatly from

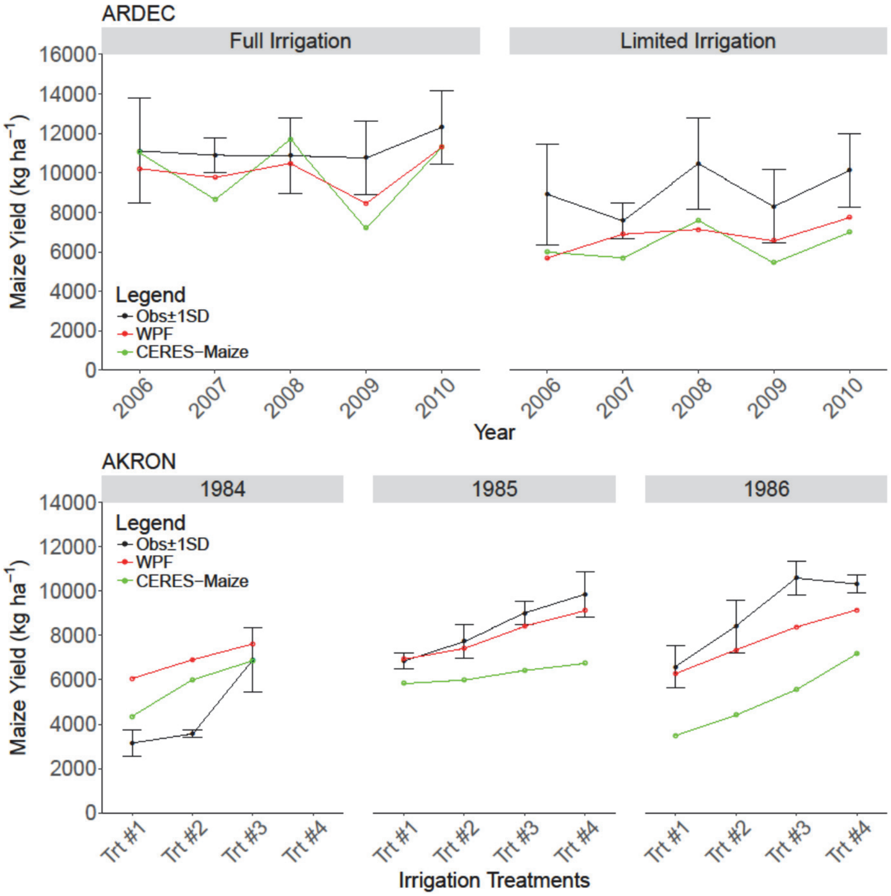

Figure 4. Predicted maize yield at the Fort Collins (ARDEC) and Akron sites in eastern Colorado using cultivar parameters (CERES-Maize) and water production function (WPF) derived from 2008-2011 data and local soil information at Fort Collins (DeJonge et al., 2011, 2012) and at Akron (Ma et al., 2003). Seasonal rainfall and irrigation were summarized from May to October for WPF prediction.

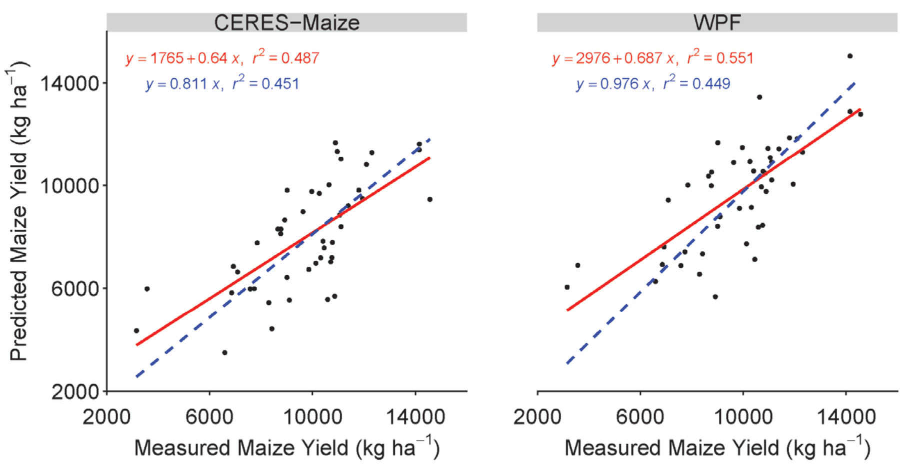

Figure 5. Measured and predicted maize yield for Greeley (2012-2013), Fort Collins (ARDEC, 2006-2010), and Akron (1984-1986) in eastern Colorado. Crop parameters (CERES-Maize) and water production function (WPF) were derived from Greeley data (2008-2011). Regression lines are shown with an intercept (equation and line in red) and with the intercept set to zero (equation and dotted line in blue).

year to year and from site to site (Nielsen et al., 2005; Nielsen and Vigil, 2017a, 2017b). Other factors, such as initial soil water content at planting, days with extreme high temperatures, and wind speed, may need to be included in regression equations to improve the predictive ability of WPFs (Nielsen and Vigil, 2017b).

Ultimately, the choice between the use of a complex crop model and a simpler WPF depends on the purpose of the user and the required accuracy. This study showed that WPFs derived from simple regression between maize grain yield or seasonal biomass and seasonal applied water (rainfall + irrigation) can be as good as, and sometimes better than, the more complex RZWQM2-CERES-Maize model in predicting maize yield or seasonal biomass. Simple WPFs would certainly be a good option if the main purpose is to estimate yield or biomass from seasonal applied water. However, complex cropping system models can provide more explanatory information (detailed soil hydrology, crop development, impacts of irrigation timing, etc.) that can give additional insights into environmental and management effects on crop growth and yield.

CONCLUSIONS

In this study, we compared RZWQM2-CERES-Maize and WPFs for predicting maize yield and seasonal biomass under various irrigation schedules at three sites (Greeley, Fort Collins, and Akron) in eastern Colorado. Results showed that WPFs were as good as RZWQM2-CERES-Maize based on r^2 value, with WPFs providing better predictions of yield in some cases based on F-test, particularly when each was calibrated to a single year of data. However, caution should be used in extrapolating the results from this study beyond the experimental sites and the model used. Further studies are needed to document the performance of system models under different soil-crop-climate conditions and to improve the simulation of biophysical processes in these complex models. More efforts are needed to overcome model limitations that hinder transferability of parameters from year to year, such as inadequate response of grain filling to variable water stress. Because water was the major limiting factor in the experiments, further comparisons of system models with multivariate regression equations is needed under both water- and nutrient-limited conditions. In addition, because different crop cultivars were used over the years across the three sites, a quantitative method for converting gene traits to cultivar parameters for the crop model could improve the performance of the model across sites. As seed companies release new maize cultivars, cultivar parameters for crop models and WPFs need to be calibrated using experimental data from the new cultivars. Improved simulation of the soil water balance in different soil types could also improve model performance in yield prediction. Ultimately, the choice between the use of a complex crop model and a simpler WPF depends on the purpose of the user and the required accuracy.

ACKNOWLEDGEMENTS

Dr. A. A. Andales's work was supported by the USDA National Institute of Food and Agriculture under Award No. 2016-68007-25066. This work was also supported by the USDA Agricultural Research Service.

REFERENCES

Abendroth, L. J., Elmore, R. W., Boyer, M. J., & Marlay, S. K. (2011). Corn growth and development. PMR 1009. Ames, IA: Iowa State University Extension.

Ahuja, L. R., & Ma, L. (2011). A synthesis of current parameterization approaches and needs for further improvements. In L. R. Ahuja & L. Ma (Eds.), *Methods of introducing system models into agricultural research* (pp. 427-440). Madison, WI: ASA, CSSA, SSSA. https://doi.org/10.2134/advagricsystmodel2.c15

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop water requirement. Irrigation and Drainage Paper No. 56. Rome, Italy: United Nations FAO.

Allen, R. G., Walter, I. A., Elliott, R. L., Howell, T. A., Itenfisu, D., Jensen, M. E., & Snyder, R. L. (2005). ASCE Standardized reference evapotranspiration equation. Reston, VA: American Society of Civil Engineers.

Araujo, L. C., Santos, P. M., Rodriguez, D., Pezzopane, J. R. M., Oliveira, P. P. A., & Cruz, P. G. (2013). Simulating Guinea grass production: Empirical and mechanistic approaches. *Agron. J., 105*(1), 61-69. https://doi.org/10.2134/agronj2012.0245

Comas, L. H., Trout, T. J., DeJonge, K. C., Zhang, H., & Gleason, S. M. (2019). Water productivity under strategic growth stage-based deficit irrigation in maize. *Agric. Water Mgmt., 212*, 433-440. https://doi.org/10.1016/j.agwat.2018.07.015

DeJonge, K. C., Andales, A. A., Ascough II, J. C., & Hansen, N. C. (2011). Modeling of full and limited irrigation scenarios for corn in a semiarid environment. *Tran. ASABE, 54*(2), 481-492. https://doi.org/10.13031/2013.36451

DeJonge, K. C., Ascough II, J. C., Andales, A. A., Hansen, N. C., Garcia, L. A., & Arabi, M. (2012). Improving evapotranspiration simulations in the CERES-Maize model under limited irrigation. *Agric. Water Mgmt., 115*, 92-103. https://doi.org/10.1016/j.agwat.2012.08.013

Doorenbos, J., & Kassam, A. H. (1986). Yield response to water. Irrigation and Drainage Paper No. 33. Rome, Italy: United Nations FAO.

Fang, Q. X., Green, T. R., Ma, L., Erskine, R. H., Malone, R. W., & Ahuja, L. R. (2010). Optimizing soil hydraulic parameters in RZWQM2 using automated calibration methods. *SSSA J., 74*(6), 1897-1913.

Fang, Q. X., Ma, L., Trout, T. J., Comas, L. H., DeJonge, K. C., Ahuja, L. R., ... Malone, R. W. (2017). Modeling N concentration and uptake for maize hybrids under growth stage-based deficit irrigations. *Trans. ASABE, 60*(6), 2067-2081. https://doi.org/10.13031/trans.12405

Getz, W. M., Marshall, C. R., Carlson, C. J., Giuggioli, L., Ryan, S. J., Romanach, S. S., ... O'Sullivan, D. (2018). Making ecological models adequate. *Ecol. Letters, 21*(2), 153-166. https://doi.org/10.1111/ele.12893

Ma, L., Ahuja, L. R., Nolan, B. T., Malone, R. W., Trout, T. J., & Qi, Z. (2012b). Root Zone Water Quality Model (RZWQM2): Model use, calibration, and validation. *Tran. ASABE, 55*(4), 1425-1446. https://doi.org/10.13031/2013.42252

Ma, L., Ahuja, L. R., Saseendran, S. A., Malone, R. W., Green, T. R., Nolan, B. T., ... Hoogenboom, G. (2011). A protocol for parameterization and calibration of RZWQM2 in field research. In L. R. Ahuja & L. Ma (Eds.), *Methods of introducing system models into agricultural research* (pp. 1-64). Madison, WI: ASA, CSSA, SSSA. https://doi.org/10.2134/advagricsystmodel2.c1

Ma, L., Hoogenboom, G., Ahuja, L. R., Ascough II, J. C., & Saseendran, S. A. (2006). Evaluation of the RZWQM-CERES-Maize hybrid model for maize production. *Agric. Syst., 87*(3), 274-295. https://doi.org/10.1016/j.agsy.2005.02.001

Ma, L., Nielsen, D. C., Ahuja, L. R., Malone, R. W., Saseendran, S. A., Rojas, K. W., ... Benjamin, J. G. (2003). Evaluation of RZWQM under varying irrigation levels in eastern Colorado. *Trans. ASAE, 46*(1), 39. https://doi.org/10.13031/2013.12547

Ma, L., Trout, T. J., Ahuja, L. R., Bausch, W. C., Saseendran, S. A., Malone, R. W., & Nielsen, D. C. (2012a). Calibrating RZWQM2 model for maize responses to deficit irrigation.

Agric. Water Mgmt., 103, 140-149. https://doi.org/10.1016/j.agwat.2011.11.005
Nielsen, D. C., & Vigil, M. F. (2017a). Defining a dryland grain sorghum production function for the central Great Plains. *Agron. J., 109*(4), 1582-1590. https://doi.org/10.2134/agronj2017.03.0131
Nielsen, D. C., & Vigil, M. F. (2017b). Water use and environmental parameters influence proso millet yield. *Field Crops Res., 212*, 34-44. https://doi.org/10.1016/j.fcr.2017.06.025
Nielsen, D. C., Saseendran, S. A., Ma, L., & Ahuja, L. R. (2012). Simulating the production potential of dryland spring canola in the central Great Plains. *Agron. J., 104*(4), 1182-1188. https://doi.org/10.2134/agronj2012.0048
Nielsen, D. C., Unger, P. W., & Miller, P. R. (2005). Efficient water use in dryland cropping systems in the Great Plains. *Agron. J., 97*(2), 364-372. https://doi.org/10.2134/agronj2005.0364
Nielsen, D. C., Vigil, M. F., & Benjamin, J. G. (2011). Evaluating decision rules for dryland rotation crop selection. *Field Crops Res., 120*(2), 254-261. https://doi.org/10.1016/j.fcr.2010.10.011
O'Grady, M. J., & O'Hare, G. M. P. (2017). Modelling the smart farm. *Info. Proc. Agric., 4*(3), 179-187. https://doi.org/10.1016/j.inpa.2017.05.001
Saseendran, S. A., Ahuja, L. R., Ma, L., Nielsen, D. C., Trout, T. J., Andales, A. A., ... Ham, J. (2014). Enhancing the water stress factors for simulation of corn in RZWQM2. *Agron. J., 106*(1), 81-94. https://doi.org/10.2134/agronj2013.0300
Saseendran, S. A., Ahuja, L. R., Ma, L., Trout, T. J., McMaster, G. S., Nielsen, D. C., ... Fang, Q. X. (2015b). Developing and normalizing average corn crop water production functions across years and locations using a system model. *Agric. Water Mgmt., 157*, 65-77. https://doi.org/10.1016/j.agwat.2014.09.002
Saseendran, S. A., Trout, T. J., Ahuja, L. R., Ma, L., McMaster, G. S., Nielsen, D. C., ... Ham, J. (2015a). Quantifying crop water stress factors from soil water measurements in a limited irrigation experiment. *Agric. Syst., 137*, 191-205. https://doi.org/10.1016/j.agsy.2014.11.005
Seidel, S. J., Palosuo, T., Thorburn, P., & Wallach, D. (2018). Towards improved calibration of crop models: Where are we now and where should we go? *European J. Agron., 94*, 25-35. https://doi.org/10.1016/j.eja.2018.01.006
Sima, N. Q., Harmel, R. D., Fang, Q. X., Ma, L., & Andales, A. A. (2018). A modified F-test for evaluating model performance by including both experimental and simulation uncertainties. *Environ. Model. Softw., 104*, 236-248. https://doi.org/10.1016/j.envsoft.2018.03.011
Trout, T. J., & Bausch, W. C. (2017). USDA-ARS Colorado maize water productivity data set. *Irrig. Sci., 35*(3), 241-249. https://doi.org/10.1007/s00271-017-0537-9
Trout, T. J., & DeJonge, K. C. (2017). Water productivity of maize in the U.S. high plains. *Irrig. Sci., 35*(3), 251-266. https://doi.org/10.1007/s00271-017-0540-1
Trout, T. J., & DeJonge, K. C. (2018). Crop water use and crop coefficients of maize in the Great Plains. *J. Irrig. Drain. Eng., 144*(6), 04018009-1-13. https://doi.org/10.1061/(ASCE)IR.1943-4774.0001309
Zhang, H., Malone, R. W., Ma, L., Ahuja, L. R., Anapalli, S. S., Marek, G. W., ... Howell, T. A. (2018a). Modeling evapotranspiration and crop growth of irrigated and non-irrigated corn in the Texas high plains using RZWQM. *Tran. ASABE, 61*(5), 1653-1666. https://doi.org/10.13031/trans.12838
Zhang, Y., Hansen, N., Trout, T., Nielsen, D., & Paustian, K. (2018b). Modeling deficit irrigation of maize with the DayCent model. *Agron. J., 110*(5), 1754-1764. https://doi.org/10.2134/agronj2017.10.0585