Crop Models Improve Growth Simulation by Including Phenological Response to Water Stress

Kyle Mankin (kyle.mankin@usda.gov)
Agricultural Research Service - Plains Area

Debora Edmunds
Agricultural Research Service - Plains Area

Gregory McMaster
Agricultural Research Service - Plains Area

Fred Fox
Agricultural Research Service - Plains Area

Larry Wagner
Agricultural Research Service - Plains Area

Timothy Green
Agricultural Research Service - Plains Area

Research Article

Keywords: Biomass, crop simulation, crop water stress, phenology, winter wheat, yield

Posted Date: September 8th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2027512/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Environmental Modeling & Assessment on November 10th, 2023. See the published version at https://doi.org/10.1007/s10666-023-09939-5.
Abstract

Crop models can provide insights into the impacts of climate and management on crop growth and yield, but most currently are limited by overly simplistic assumptions about phenological development and response to water stress. We assessed winter wheat growth and yield performance of three crop models with lineage to the EPIC crop submodel. SWAT adopted the EPIC approach with few modifications, WEPS added new biomass accumulation, partitioning, and canopy approaches linked to key phenological development stages, and UPGM added to WEPS a detailed phenology component simulating responses to water stress. The models were evaluated with default parameters and compared to experimental data for winter wheat (Triticum aestivum L.) from two sites and a range of water-stress conditions for yield, aboveground biomass, biomass partitioning, canopy height, harvest index, and leaf area index. All models simulated yield very well (index of agreement $d \geq 0.93$), but differences in model performance were increasingly evident for biomass ($d = 0.91$ [WEPS] to 0.86 [SWAT]), final canopy height ($d = 0.68$ [UPGM] to 0.44 [SWAT]), and harvest index ($d = 0.61$ [WEPS] to 0.43 [SWAT]). Errors in biomass simulation were most evident in the grain-filling period late in the growing season. Both WEPS and UPGM exhibited improved simulation of biomass and other response variables by including more explicit simulation of phenological response to water stress. The consistent improvement in winter wheat growth and yield simulation achieved with detailed phenology simulation provides an incentive to develop and test detailed phenology simulation components for other crops: currently 11 crops are simulated in UPGM, although the phenological parameters are uncalibrated. Better modeling linkages of water-stressed phenological development with other physiological processes will be critical to inform crop production where water stress and irrigation limitation are concerns.

1. Introduction

Crop phenology strongly influences crop yield, biomass accumulation, and other plant processes, such as the magnitude and timing of evapotranspiration (ET), in agricultural ecosystems. Crop yield is a key element of farm profit. ET is the primary hydrologic outflow from croplands and controls the fraction of precipitation and irrigation inflows that is available to contribute to soil water storage, streamflow, aquifer recharge, and ecosystem function [1-2]. Crop growth and development interacts with management, water availability, and local conditions to influence yields and hydrologic responses [3]. Accurate representation of crop processes is essential to hydrologic simulation in watershed models. Yet watershed models typically use simplistic assumptions about crop growth and development and their influence on crop yield, biomass accumulation, and other plant process, such as ET.

The Environmental Policy Integrated Climate (EPIC [4-5]) includes a crop growth submodel [6] that has been applied widely and modified in other models to meet specific modeling objectives including Soil Water Assessment Tool (SWAT [7-8]), Wind Erosion Prediction System (WEPS [9]), and Unified Plant Growth Model (UPGM [10]). Important processes of the EPIC crop growth submodel are described in the Background Information.
McMaster et al. [11] assessed performance of the three crop models in simulating the timing of five winter-wheat developmental stages under various levels of water stress. Start of senescence (S) and physiological maturity (M) were compared among SWAT, WEPS and UPGM. Jointing (J) was compared between WEPS and UPGM. Heading (H) and start of anthesis (AS) were assessed for UPGM alone. UPGM was the only model to explicitly simulate all five developmental stages (J, S, H, AS, M). They found that UPGM simulated J, S, and M as well or better than WEPS and always better than SWAT. Unanswered by McMaster et al. [11] is the question of whether the demonstrated superiority in simulating phenological timing translates into improved simulation of crop yield and biomass growth components across varying environments. The objective of the present study is to evaluate whether the improved phenology model in UPGM also improves the simulation of other important winter wheat growth and development processes (i.e., yield, aboveground biomass, canopy height, harvest index, and leaf area index [LAI]) under various levels of water stress compared to SWAT and WEPS. Importantly, all three crop models are executed within the Agricultural Ecosystems Services (AgES [12]) hydrology model, thereby using a common platform for other system processes that drive crop growth and development.

2. Background Information

The EPIC crop growth submodel is the basis for the three models studied here. In EPIC, biomass is accumulated daily as a product of radiation use efficiency (RUE) and intercepted photosynthetically active radiation, which is calculated based on simulated LAI. Daily biomass is partitioned among above- and below-ground fractions based on a ratio that changes from 40% (at germination) to 20% (at maturity). Canopy height is also calculated following a crop-specific LAI-based curve. LAI is accumulated daily according to heat units (HU) and is attenuated by crop stress. Biomass is translated to crop yield as a fraction (harvest index, HI) of aboveground biomass. HI is assumed to be maximized by breeding for a given cultivar and relatively stable across a range of environments [13]. A water-stress factor (Wstrs) is calculated daily based on a ratio of actual water transpiration ($T_a$) to potential water transpiration ($T_p$) as,

$$\text{Wstrs} = \frac{T_a}{T_p} \tag{1}$$

In this formulation, there is no water stress for $\text{Wstrs} = 1$, and the crop is fully water-stressed when $\text{Wstrs} = 0$. Subsequently, Wstrs affects both daily biomass accumulation and HI, where HI is influenced by water stress during crop stages between 0.3 and 0.9 of maturity, as measured by a heat-unit index (HUI), with maximum sensitivity at HUI of 0.6, near crop anthesis. The HUI on day $i$ after planting (Eq. 2) is calculated as a fraction of accumulated heat units (HU, Eq. 3) to theoretical HUs from planting to maturity (THUM):

$$\text{HUI}_i = \frac{\sum_{j=1}^{i} HU_j}{THUM} \tag{2}$$

$$HU_i = T_{avg,i} - T_{base}, \text{ for } T_{avg,i} \geq T_{base}$$

$$HU_i = 0, \text{ for } T_{avg,i} < T_{base} \tag{3}$$
where HUI is heat unit index (fraction: 0 at planting, 1 at maturity) on day i after planting, HU\textsubscript{i} is heat units accumulated on day i (°C day, or growing degree-day [GDD]), THUM is theoretical HUs accumulated at physiological maturity (°C day), T\textsubscript{avg} is average daily air temperature (°C), and T\textsubscript{base} is crop-specific base temperature (°C). Phenological development in EPIC follows the common approach of accumulating daily HUs (Eq. 3). Only a few developmental events are specified. The crop-specific input value of HU required to reach maturity (THUM) is fixed for each crop. The HUI is used to specify dates of leaf-area growth (different rate equations before and after HUI of crop-specific start of leaf decline), root depth (maximum when HUI reaches 0.4), economic yield (nonlinear HI increase from HUI of 0.5 to 0.9), and physiological maturity (HUI of 1). A couple potential problems with this formulation in Eq. 2 are that delayed seedling emergence and/or vernalization can result in significant errors in estimating maturity and are not reflected in the pre-determined HUI curves. Further, use of these fixed values restricts model ability to simulate the influence of water stress on phenology. Although water stress in EPIC influences biomass accumulation, crop yield, crop height, and LAI accumulation, which in turn affects water use, the HUs required to progress between developmental stages are fixed and independent of water stress. Therefore, while water stress affects the magnitude of crop growth processes, it does not reflect the important changes in timing of crop development.

Of the three models (SWAT, WEPS, and UPGM) evaluated in this study, the SWAT version 2009 [13] adopted the EPIC crop growth submodel with few modifications of the above processes. Each model provides a different approach to simulating phenology, as described by McMaster et al. [10–11], based on the thermal-time approach to simulating phenological development (Eqs. 2, 3). SWAT uses the approach described above, and the phenological approaches used in WEPS and UPGM will be described briefly herein, along with other modifications.

WEPS made some significant modifications and additions to the EPIC plant growth component to increase the accuracy and robustness of simulating the canopy important to simulating wind erosion. WEPS also uses the EPIC crop growth submodel with static values of phenological stages that do not respond to water stress. WEPS also defines the HUI beginning at planting (P) but defines a HU-based curve for reproductive partitioning. We associate the double-ridge (DR) event with the beginning of reproductive partitioning. Jointing (J) is assumed to occur 180 GDDs after DR [11, 14–15]. Unlike EPIC, in which LAI is increased on a HU-based curve, WEPS accumulates and partitions biomass to various plant parts and then calculates the resulting LAI from leaf biomass using a crop-specific leaf area conversion factor [9]. This allows WEPS to develop LAI uniquely each growing season based on actual conditions, including water stress, rather than an assumed curve. WEPS uses an S-shaped curve to partition aboveground biomass accumulation to leaf, stem, and reproductive components based on HU (Eqs. 2, 3). As such, WEPS uses an HU-based approach, rather than EPIC’s HI-based approach, to simulate yield [10]. HI is then calculated from total aboveground biomass and yield. The curve used for crop height increase is independent of LAI. It is a separately specified curve tied to GDD (stage of growth) and the incremental potential increase applied on each day is reduced by stress level (maximum of water and temperature stress) on that day. These WEPS modifications to crop growth make it more mechanistic in many processes, including determining plant height, partitioning biomass to plant components rather than dividing them up after the fact, and yield.
The underlying rationale for developing UPGM was to combine the various modifications to the EPIC plant growth submodel into one unified crop model that would facilitate parameterization and future model development, while maintaining important differences in addressing slightly differing model objectives. The WEPS plant growth submodel was used as the foundation for UPGM because the underlying structure was the most modified of the various models and therefore made it easier to incorporate other model improvements. A more mechanistic simulation approach was desired for UPGM model development, beginning with a more detailed phenology submodel that responded to varying water deficits that could then be used in defining the trajectory of biomass, yield, canopy height, etc. It also would allow for adjusting the timing of the aboveground biomass curves. UPGM integrated the WEPS plant growth component [10] with a phenological development component from PhenologyMMS [16]. UPGM explicitly defines the thermal time in HUs (GDDs) required to reach many important developmental events for both non-stressed (GN) and water-stressed (GS) growth conditions (Table 1). GN and GS parameters for each crop and developmental stage were taken from PhenologyMMS [15–16]. UPGM adjusts the HUs required to reach the next developmental event using a piecewise linear interpolation between GN and GS within a range of 0.4 < Wstrs < 0.8 (see Fig. 1a for a schematic representation). Water stress can accelerate development (when GS < GN), slow development (when GS > GN), or have no affect (when GS = GN), depending on the crop and growth stage. Figure 1b shows quantitative functions of thermal time versus Wstrs for eight different developmental stages of winter wheat using the values given in Table 1. Knowing the HUs when a growth stage was reached would then allow this timing to be used rather than a set input parameter that did not respond to water stress.

### Table 1
UPGM thermal times (growing degree-days, GDD, 0°C T<sub>base</sub>) between developmental events for winter wheat (*Triticum aestivum* L.) under non-stressed (GN) and water-stressed (GS) conditions

(adapted from 11, 15).

<table>
<thead>
<tr>
<th>Phenology Parameter</th>
<th>GDD between Winter Wheat Developmental Events</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E→TI</td>
<td>1/1→</td>
</tr>
<tr>
<td>GN</td>
<td>200</td>
<td>180</td>
</tr>
<tr>
<td>GS</td>
<td>200</td>
<td>180</td>
</tr>
</tbody>
</table>

E = emergence, TI = tiller initiation, 1/1 = January 1, SR = single ridge, DR = double ridge, IES = internode elongation starts, TS = terminal spikelet, J = jointing, FLC = flag leaf complete, B = booting, H = heading, AS = anthesis starts, M = physiological maturity, HR = harvest ready.

### 3. Materials And Methods

SWAT, WEPS, and UPGM were incorporated into the USDA-ARS Agricultural Ecosystems Services (AgES) hydrologic model [12, 17]. AgES is a spatially distributed hydrologic/water quality model that adapted hydrologic components from the European J2K model [18], nutrient components from the SWAT model [13], and crop growth components from either SWAT, WEPS, or UPGM in an object modeling framework [19].
AgES hydrologic components have been tested extensively [20–22] and have provided a consistent basis for applying climate, water stress, and other site factors to the three models under comparison.

When applying these three crop models within a watershed model with many fields and management units, modelers typically use the default crop parameters without any site-specific calibration. This is the main reason we did not optimize crop parameters for this study. For each test site below, other parameters in AgES that control soil hydrology and ET were calibrated previously using “global” and soil layer parameters [11–12].

3.1 Experimental Comparison Data

Two experimental datasets growing winter wheat (\textit{Triticum aestivum} L.) in semi-arid northeastern Colorado were used. Site details are fully described elsewhere as noted below, and only key elements briefly summarized herein. Both data sets measured five developmental events: jointing (J; when the first node is above the soil surface), start of senescence (S; flag leaf blade growth completed as measured by formation of the ligule on the flag leaf, which is the end of leaf appearance on a shoot), heading (H; appearance of spike emerging from the flag leaf, not including awns), start of anthesis (AS; when the first anthers are visible in the spike), and physiological maturity (M; denoted by the absence of all green color in the spike and subtending peduncle) ([23] as described in [11]).

The Greeley dataset [11] was a 3-year study (Sep 2008 to Aug 2011) of various height and maturity classes of hard red and white winter wheat at the USDA-ARS Limited Irrigation Research Farm (LIRF) near Greeley, CO (40° 26’ 57.13” N; 104° 38’ 12.04” W; 1427 masl). Soils were deep and well drained fine sandy loam to clay loam. Weather data were available on-site (CoAgMet Station GLY04; Table 2). The full study was comprised of a split-plot design: 24 wheat varieties randomized within 5 irrigation treatments plots with 3 replications (360 plots, 7 m² each). The mean of the 18 medium-maturity varieties (Above, Ankor, Avalanche, Baca, Bill Brown, BondCL, CO940610, Danby, Goodstreak, Hatcher, Jagalene, Keota, NuDakota, Platte, Prairie Red, RonL, Sandy, Yuma) and two of the irrigation treatments (fully irrigated [Wet], dryland [Dry]) were used in this modeling study. On fully irrigated plots (Wet), surface drip irrigation was applied to minimize water stress (maintaining soil water > 75%); dryland plots (Dry) received no irrigation. Fall planting operations consisted of disking (10-cm depth), fertilizer application (per soil test), roller harrowing, planting (2 Oct 2008, 6 Oct 2009, 8 Oct 2010), and portable sprinkler irrigation (24, 25, and 43 mm of water, respectively) to ensure emergence. Phenology data were typically collected 3 d per week. Dates on which the first shoot and half the shoots within a plot reached a developmental stage were recorded. Other plant measurements also were combined for the 18 medium-maturity varieties. Total aboveground biomass (and leaf blades, stem + leaf sheaths, inflorescence fractions if present) was measured each year at pre-jointing (early spring when still in the rosette form), jointing, anthesis, mid-grain filling (not in 2008-09), and maturity. All samplings before maturity harvested 0.5 m row lengths whereas maturity harvested 1 m row lengths. A plot combine was used to harvest the entire plot for yield and determine percent moisture. Harvest index (HI) was determined from the maturity aboveground biomass sampling. Canopy height was also generally measured each year at pre-jointing (early spring when still in the rosette form), jointing, anthesis, mid-grain filling, and maturity, with 4 to 9 randomly selected points in the plot.
The Drake site [11, 24] was a 4-year dryland study (harvest years 2004 to 2007) on a 109-ha farm field (Drake Farm) east of Fort Collins, CO near Ault, CO (40.61°N, 104.84°W, 1559–1585 masl). Soils were Wagonwheel coarse silty-loam, Colby fine silty-loam, and Kim fine sandy-loam with 0.65 to 8.4% slopes. Weather data were available onsite ([24]; Table 2). Alternating strips (~ 120 m wide) were in 2-year wheat-fallow rotation. A medium-maturity winter wheat variety (Above) was planted in alternating strips (1 Oct 2003, 23–27 Sep 2004, 1 Oct 2005, 3 Nov 2006). Phenology data were typically collected 3 d per week as a repeated measurement of tagged plants from ten 30 x 30 m sites split between two of the strips (2 subsamples per site, 10 plants per subsample). Dates on which each shoot reached a developmental stage were recorded. Canopy height was measured for each tagged plant. Total aboveground biomass was sampled by hand within the phenological plot areas. Final grain yield was based on yield monitor data at 10-m resolution within each sample area.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Greeley Site</th>
<th>Drake Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total P, GS, 1 Sep–30 Jun (mm)</td>
<td>234</td>
<td>195</td>
</tr>
<tr>
<td>Full Irr, GS, 1 Sep–30 Jun (mm)</td>
<td>257</td>
<td>374</td>
</tr>
<tr>
<td>P + Irr, GF, 1 May–30 Jun (mm)</td>
<td>273</td>
<td>337</td>
</tr>
<tr>
<td>Mean T, GS, 1 Sep–30 Jun (°C)</td>
<td>6.7</td>
<td>6.9</td>
</tr>
<tr>
<td>Mean T, GF, 1 May–30 Jun (°C)</td>
<td>16.4</td>
<td>15.6</td>
</tr>
</tbody>
</table>

LTM = long-term mean (2002-18), P = precipitation, Irr = irrigation, T = temperature, GS = growing season, GF = grain filling period.

### 3.2 Model Runs and Evaluation

AgES models were developed for each site using onsite weather data. The Greeley model was set up to run the two treatments (Wet, Dry) as independent hydrologic response units (HRUs) over the 3-year study (2009, 2010, 2011 named by year of harvest). The Drake model was simulated as a 56-ha watershed with 27 HRUs; two summit locations, HRU 23 (2004, 2006) and HRU 24 (2005, 2007), were used for this study to minimize run-on effects and help ensure water limitation. AgES was run using each of the three crop models with default crop parameters to focus on the primary objective of comparing performance of three similar (EPIC-based) crop models with enhanced phenology (WEPS, UPGM) and water-stress response (UPGM) without the influence of calibration skill. Models were assessed using several statistical performance measures: root-mean square error (RMSE; [25]), relative error (RE = [predicted mean – observed mean] / observed mean, %), index of agreement (d; [25]), and normalized objective function (NOF = RMSE / observed mean).
4. Results And Discussion

All three crop models performed well in simulating winter wheat yield (Table 3, Fig. 2). Across all sites, years, and soil water conditions, index of agreement (d) exceeded 0.9 for all models (d = 0.97 [SWAT], 0.96 [WEPS], 0.93 [UPGM]). Normalized objective function (NOF) demonstrated slightly decreasing performance from SWAT (0.16) to WEPS (0.23) to UPGM (0.32). WEPS and UPGM generally overpredicted yield (positive RE = 10.5% [WEPS], 19.6% [UPGM]) whereas SWAT slightly underpredicted yield (negative RE = -3.4%), particularly for Greeley Wet. Observed yield for Greeley Dry was similar over the 3 years (3270–3762 kg/ha), but all models simulated greater yield differences (SWAT: 3073–4046; WEPS: 2883–5159; UPGM: 3271–5631 kg/ha). Even without calibration, all models demonstrated skill in simulating wheat yield across the full range of water conditions spanning 10 crop-years. However, no clear benefit was evident for the HU approach to simulating yield used by WEPS and UPGM over the HI approach used by SWAT.

Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>SWAT</th>
<th>WEPS</th>
<th>UPGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE (%)</td>
<td>d</td>
<td>NOF</td>
<td>RE (%)</td>
</tr>
<tr>
<td>Yield</td>
<td>-3.4</td>
<td>0.97</td>
<td>0.16</td>
</tr>
<tr>
<td>AG biomass</td>
<td>-20.2</td>
<td>0.86</td>
<td>0.35</td>
</tr>
<tr>
<td>Canopy height</td>
<td>49.4</td>
<td>0.44</td>
<td>0.59</td>
</tr>
<tr>
<td>Harvest index</td>
<td>37.7</td>
<td>0.43</td>
<td>0.48</td>
</tr>
</tbody>
</table>

RE = relative error (%), d = index of agreement, NOF = normalized objective function, AG = aboveground.

All three crop models performed well in simulating final aboveground biomass (Table 3, Fig. 3). Performance in simulating final biomass declined slightly compared to yield performance for all models, but d exceeded 0.86 and NOF was less than 0.35 for all models. WEPS (RE = -11.6%, d = 0.91, NOF = 0.30) performed best, followed by UPGM (-13.7%, 0.87, 0.33) and SWAT (-20.2%, 0.86, 0.35). For biomass simulation, WEPS and UPGM, which used an HU partitioning approach outperformed SWAT with its standardized LAI-curve approach. This may point to a benefit in accumulating HUs as attenuated by water stress in simulating crop biomass.

Multiple in-season observations at the Greeley site demonstrated that WEPS and UPGM closely simulated aboveground biomass throughout the growing season, whereas SWAT results were consistently lower and underpredicted observed values (Fig. 4). In cases where errors developed in biomass simulation, they generally occurred late in the season. Final biomass was substantially underpredicted by all models for two site years: Greeley Wet 2009 (Figs. 3, 4) and Drake 2006 (Figs. 3, 5). For Greeley Wet 2009, in-season data shows that all models performed reasonably well up to day 149, but simulated biomass plateaued after about day 180 while observed biomass increased up to day 199 (Fig. 4a). Other treatment-years at Greeley showed an end-of-season plateau in observed biomass that was captured by simulated results. This may
indicate that the models are not capturing a periodic climate or management influence on biomass accumulation. For example, the underpredicted Greeley 2009 and Drake 2006 site years were the warmest and driest years during the grain-filling period (1 May to 30 June) at their respective sites during the study period (Table 2). In some cases, UPGM better tracked early-season while WEPS better captured end-of-season biomass accumulation (Greeley Wet 2011, Drake 2004) or vice versa, with WEPS better tracking early-season biomass accumulation (Greeley Wet 2010, Greeley Dry 2011), which suggests the need to calibrate timing of phenological development stages with biomass accumulation (Figs. 4, 5).

SWAT simulates daily accumulated total biomass; WEPS and UPGM also simulate daily biomass partitioned to leaves, stems, straw (leaves + stems), and seed heads (Fig. 6) as well as roots (not assessed in this study). SWAT consistently underpredicted total biomass compared to WEPS and UPGM except for three occasions: UPGM Greeley Wet 2009 (Fig. 6a) and WEPS Greeley Dry 2011 (Fig. 6f) at the very end of the season, and WEPS Drake 2004 HRU 23 (Fig. 6g) for most of the season.

The considerable effect of water stress on simulated timing of developmental stages by WEPS and UPGM is also evident in Fig. 6, as represented by the five vertical lines shown in each diagram for the simulated date of J, S, H, AS, and M stages. WEPS and UPGM differ in the magnitude and timing of biomass partitioning leading to the differences between the biomass accumulation curves. Biomass data collected near day 150 for 8 of the 10 site years often were well simulated for seed heads (WEPS, except Fig. 6c; and UPGM, except Fig. 6c,i), straw (WEPS, except Fig. 6a,b,g; and UPGM, except Fig. 6a,b,e), and total biomass (WEPS, except Fig. 6g; and UPGM, except Fig. 6a,b,c,d,e,i).

The wheat canopy was consistently simulated by SWAT to achieve maximum height regardless of site and climate factors, whereas UPGM and WEPS better reflected observed heights (Fig. 7). Simulated final canopy height increased in rough agreement with observed heights (d = 0.60 [WEPS], 0.68 [UPGM]), although height was underpredicted by both WEPS (RE = -42%) and, to a lesser extent, UPGM (-34%) (Table 3, Fig. 7). The SWAT method using a HU-adjusted fraction of maximum canopy height was accurate only for non-limiting conditions, such as the fully irrigated Greeley Wet conditions of 2009 and 2011 (Figs. 7, 8). Overall, SWAT significantly overpredicted height (RE = 49%) (Table 3, Fig. 7c). Multiple in-season observations at the Greeley site demonstrated that all models simulated a mid-season increase followed by late-season plateau (Fig. 8). Although WEPS and UPGM allow simulation of season-specific canopy height development, both models have difficulty capturing both timing and magnitude of the growth without further adjustments to parameters.

By design, HI in SWAT is relatively stable (here, about 0.50) across environmental conditions, while HI in WEPS and UPGM better predicts the observed values (Fig. 9). HI ranged from 0.23 to 0.60 across sites and years, so WEPS and UPGM, which can account for this variability, have a clear advantage. All models generally overpredicted HI. WEPS (RE = 25%, d = 0.61, NOF = 0.36) performed best, followed by UPGM (35%, 0.54, 0.47) and SWAT (38%, 0.43, 0.48) (Table 3, Fig. 9). Interestingly, SWAT had the greatest overprediction of HI (RE = 38%), which combined with the greatest underprediction of aboveground biomass (-20%) and resulted in the most accurate prediction of yield (-3.4%). WEPS and UPGM do not use HI in calculation of yield, so a similar comparison is not relevant.
No LAI observations were made, so model performance can only be assessed by comparison among models (Fig. 10). For SWAT, the timing varied slightly among plots and years, but the pattern was similar with a consistent peak LAI of 3.0, because SWAT was not responding to variable water conditions. The magnitude and timing of LAI varied only slightly among site years for SWAT but varied considerably for WEPS and UPGM in response to site conditions. In many cases, WEPS and UPGM had greater peak LAI than SWAT, sometimes exceeding a reasonable maximum LAI of 5. Examining leaf senescence after peak LAI, all three models overpredicted final LAI.

5. Conclusions

Three crop models with lineage to the EPIC crop growth submodel were compared to experimental winter wheat (*Triticum aestivum* L.) growth and yield data. Complexity in representing crop phenology and response to water stress increased from SWAT to WEPS to UPGM, and our hypothesis was that this increasing complexity would result in increasingly accurate simulation of all crop growth variables across a wide range of environmental conditions. This hypothesis was generally confirmed, particularly when considering most of the simulated processes were very similar in WEPS and UPGM. All crop models performed well in simulating yield (index of agreement \([d] \geq 0.93\)). WEPS and UPGM outperformed SWAT for other growth variables, including biomass, canopy height, and harvest index.

These results show benefits from more explicit simulation of phenology and response to water stress in crop models. We expect that the framework provided by explicit phenological simulation will allow further improvements in model performance that are not possible without the ability to simulate development-specific differences in biomass partitioning, leaf-area expansion, and grain filling. Addition of improved water-stress response will provide better simulation of future, non-stationary climate impacts on crop growth and development, which will inform farm to regional crop management and planning decisions. Such enhancements are also expected to improve simulation of other crop processes, such as evapotranspiration and biomass residue production and management.

This study focused on winter wheat, an important crop regionally and globally in semi-arid climates, where irrigation supply is limited and water stress is common during the growing season. Further development is needed to expand phenological parameterization to other crops, particularly those grown in water-limited production systems.

Declarations

Acknowledgements

The USDA is an equal opportunity provider and employer. Nathan Lighthart, Holm Kipka and Roger Marquez provided Java programming support for integration of crop submodels into AgES and supported AgES simulations. Rob Erskine and Gerald Buchleiter assisted with data collection and processing at the project sites. David Drake farmed winter wheat on the Scott Field and provided access for data collection.

Ethical approval. Not applicable, as this research was not conducted on human and/or animal subjects.
Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors’ contributions. Debora Edmunds, Gregory McMaster, and Timothy Green contributed to the study conception and design. Data collection and analysis were performed by Kyle Mankin, Debora Edmunds, Gregory McMaster, and Timothy Green. Modeling was conducted by Debora Edmunds, Fred Fox, Larry Wagner, and Timothy Green The first draft of the manuscript was written by Kyle Mankin, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding. This project was conducted using federal funds appropriated to the Agricultural Research Service for research under project number 3012-13660-009-00D.

Availability of data and materials. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References


**Figures**

(a) Conceptual diagram of thermal time to complete a developmental event with linear interpolation from non-stressed (GN, Wstr=1) to water-stressed (GS, Wstr=0) phenological parameters, where water stress can accelerate (decrease thermal time) or delay (increase thermal time) development. (b) Example of actual developmental stages for winter wheat showing 8 of 10 stages from Table 1 (E→TI (GS=GN) and AS→M (GS<GN) are not shown for graphical purposes only).
Figure 2

Yield performance.
Figure 3

Final aboveground (AG) biomass performance.
Figure 4

Aboveground biomass simulated time series and observed data for the Greeley site.
Figure 5

Aboveground biomass simulated time series and observed data for the Drake site.
Figure 6

UPGM and WEPS simulated biomass partitioning (total aboveground biomass, leaves, stems, straw [leaves + stems], seed heads) compared to SWAT (total only) and observed data for Greeley Wet (a: 2009, b: 2010, c: 2011), Greeley Dry (d: 2009, e: 2010, f: 2011), Drake HRU 23 (g: 2004, h: 2006), and Drake HRU 24 (i: 2005, j: 2007). Also shown are vertical lines for five observed phenological stages: jointing (J), start of senescence (S), heading (H), anthesis starts (AS), maturity (M).
Figure 7

Final canopy height (ht.) performance.
Figure 8

Canopy height simulated time series and observed data for the Greeley site.
Figure 9

Harvest index performance.
Figure 10

Leaf area index (LAI) simulated time series comparison for each site and year.