

Original papers

A parsimonious Bayesian crop growth model for water-limited winter wheat

Pratishtha Poudel^{a,1}, Phillip D. Alderman^{a,*}, Tyson E. Ochsner^a, Romulo P. Lollato^b^a Oklahoma State University, Department of Plant and Soil Sciences, 371 Agricultural Hall, Stillwater, OK, 74078, United States of America^b Kansas State University, Department of Agronomy, 2709 Throckmorton Plant Sciences Center, 1712 Claflin Road, Manhattan, KS, 66506, United States of America

ARTICLE INFO

Keywords:

Crop growth modeling
Bayesian analysis
Parameter estimation
Water limitation
Winter wheat

ABSTRACT

Dynamic crop models are widely used to simulate crop production, but are often complex and thus face parameter non-identifiability issues. In this study, we demonstrate a simple dynamic model for crop growth within a Bayesian hierarchical framework and quantify improvements in predicted crop growth patterns by inclusion of a dynamic water balance. Seven crop parameters and four water balance parameters were estimated by the model from data on leaf area index, biomass, yield, and plant available water over the growing season across multiple environments. Posterior median values for Willmott agreement index (d) and Nash–Sutcliffe efficiency (NSE) showed that the model predicted leaf area index (LAI; $d=0.89$; $NSE=0.62$) and biomass well ($d=0.98$; $NSE=0.92$) with less success for plant available water (PAW; $d=0.75$; $NSE=-0.03$) and grain yield ($d=0.90$; $NSE=0.40$). Inclusion of a water balance component raised median NSE from 0.57 to 0.62 for LAI, 0.74 to 0.92 for biomass, and -0.58 to 0.40 for grain yield. The median and highest density interval (HDI) were biologically plausible for the parameters canopy light extinction coefficient (median=0.46; HDI=[0.38, 0.55]), maximum leaf area index (median=6.66; HDI=[6.06, 7.24]), and crop evapotranspiration coefficients for the initial (median=0.03; HDI=[0.000004, 0.08]) and mid-season (median=0.32; HDI=[0.23, 0.42]) growth stages. Posterior values for the potential radiation use efficiency (RUE_p ; $g\ MJ^{-1}\ (^{\circ}C-d)^{-1}$) parameter (median=2.17; HDI=[1.99, 2.34]) were not biologically plausible indicating a need for model improvement.

1. Introduction

Crop models are commonly used to simulate growth and yield and to understand the mechanistic interplay between crops and environments (Attia et al., 2016; Asseng et al., 2014). Successful usage of these models require a thorough understanding of the biological system, along with accurate parameter estimation, calibration, and validation. However, complete understanding of all underlying mechanistic processes is difficult to achieve in complex systems such as cropping systems. In addition, specifying ordinary differential equations (ODEs) for numerous processes in order to mechanistically represent the full system results in a level of model complexity that frequently renders such models non-identifiable due to the dearth of sufficiently informative data to fit them (Lamsal et al., 2018). Standard crop growth models such as DSSAT-CSM (Jones et al., 2003), APSIM (Keating et al., 2003), WOFOST (Van Diepen et al., 1989) and CropSyst (Stöckle et al., 2003) have a large number of parameters and input data requirements that are beyond the scope of many field-based studies. As a result, a more parsimonious approach that incorporates a minimum level of detail

required to represent the system under study may be preferable in many situations. In this study, we propose such a parsimonious approach to crop growth modeling by introducing a simple mechanistic ODE model into a Bayesian hierarchical framework.

The approach we propose can accommodate parameter uncertainty and random noise in the cropping system probabilistically. In so doing, we simultaneously account for the incomplete knowledge of the system and non-systematic sources of variability in a way that standard crop growth models cannot. This type of hybrid approach is gaining attention in other scientific areas such as food science (Yang et al., 2021) and medicine (Mascheroni et al., 2021). While it is important to specify the fundamental processes in a manner consistent with a mechanistic understanding of the system dynamics, accounting for random noise with a stochastic model component allows for modeling sources of variation not explicitly handled by the deterministic components of the model. Along with accounting for random noise, the Bayesian approach also allows us to inform the parameter estimation process with prior knowledge about the system and characterize the full joint distribution of parameters of interest, including their associated

* Corresponding author.

E-mail address: phillip.alderman@okstate.edu (P.D. Alderman).¹ Present address: Purdue University, Department of Agronomy, 915 Mitch Daniels Blvd, West Lafayette IN, 47907.

uncertainties, rather than point estimates alone. This approach provides a more holistic perspective on how individual parameters which govern key processes relate to each other to affect the evolution of the system as a whole over the growing season.

This study is conducted in the southern Great Plains (SGP) of the United States. Winter wheat (*Triticum aestivum* L.) is an important crop grown in this region and soil water balance is one of the important drivers of yield variability (Munaro et al., 2020). Precipitation patterns vary both temporally and spatially which results in diverse environments across the region (Maulana et al., 2019). It is thus reasonable to consider modeling soil water balance as a component within a crop growth model for this region even within a minimalist crop growth model. Crop models accounting for soil water effects on crop growth range from relatively simple models, such as AquaCrop (Steduto et al., 2009) to more complex models such as DSSAT-CERES (Jones et al., 2003). Although varying in complexity, they are fundamentally similar in simulating water balance as a function of a drained upper limit (field capacity) and a lower limit (permanent wilting point). Most of these models simulate soil water balance for multiple layers within a soil profile, although some models simulate the soil water in the rooting zone as a single layer (Steduto et al., 2009). As recently reviewed by Tenreiro et al. (2020), the methods for simulating infiltration in crop models range from the relatively simple runoff curve number method (United States Department of Agriculture, Soil Conservation Service, Engineering Division, 1986) to more mechanistic approaches based on the Richard's equation (Richards, 1931). For simulating evapotranspiration, many crop models use an approach derived from the Penman–Monteith equation (Penman, 1948; Monteith, 1965), while some also provide implementations of the Priestley–Taylor (Priestley and Taylor, 1972) or Hargreaves methods (Hargreaves and Samani, 1982).

These types of dynamic simulation models are frequently used to explain genotype by environment interaction for multiple genotypes across years. With extensive datasets from multi-environment experimental trials, it becomes particularly important to account for random noise or non-systematic sources of variability in a system. Poudel et al. (2022) presented initial work in the integration of mechanistic and random components to model repeated measures data on LAI and biomass over time for irrigated and rainfed winter wheat in the SGP. The specific objective of this study is to present a new, re-conceptualized model to better represent the underlying processes related to vernalization, plant available water (PAW) and grain yield.

2. Methodology

2.1. The ODE crop growth model

Table 1 presents a description of the state variables, parameters, and input variables in the ODE model. The table is followed by description of the equations.

2.1.1. Thermal time

The development of the crop is expressed the function of temperature over time as well as the vernalization status of the crop. The rate of change of thermal time is specified as:

$$\frac{dT T_t}{dt} = f_{TT\ t} \cdot f_{vrm\ t} \quad (1)$$

where,

$\frac{dT T_t}{dt}$ is the rate of change of thermal time at time t .

$f_{TT\ t}$ is the thermal time factor, calculated as shown below.

$f_{vrm\ t}$ is factor corresponding to the vernalization requirement, calculated as shown below.

Table 1

Description of the state variables, inputs, and parameters in the ODE model.

Notation	Units	Type	Description
$T T_t$	°C-d	State variable	Cumulative thermal time at time t
$V R N_t$	d	State variable	Cumulative vernalization at time t
$L A I_t$	m ² m ⁻²	State variable	Leaf area index at time t
$B M_t$	g m ⁻²	State variable	Biomass at time t
$Y L D_t$	g m ⁻²	State variable	Grain mass at time t
$P A W_t$	mm	State variable	Plant available water content at time t
α	(°C-d) ⁻¹	Parameter	Relative rate of LAI increase
$T T L$	°C-d	Parameter	Thermal time to the end of leaf expansion
$R U E_p$	g MJ ⁻¹ (°C-d) ⁻¹	Parameter	Potential radiation use efficiency
K	dimensionless	Parameter	Light extinction coefficient
$L A I M A X$	m ² m ⁻²	Parameter	Maximum leaf area index
$senrate$	(°C-d) ⁻¹	Parameter	Relative rate of senescence
$vreq$	d	Parameter	Vernalization requirement
$wsen$	dimensionless	Parameter	Water stress sensitivity
$K_{c\ ini}$	dimensionless	Parameter	Early season crop coefficient
$K_{c\ mid}$	dimensionless	Parameter	Mid season crop coefficient
$K_{c\ end}$	dimensionless	Parameter	Late season crop coefficient
$T avg_t$	°C d ⁻¹	Model input	Observed daily temperature at time t
$S R A D_t$	MJ m ⁻² d ⁻¹	Model input	Observed daily solar radiation at time t
P_t	mm d ⁻¹	Model input	Observed daily precipitation at time t
I_t	mm d ⁻¹	Model input	Irrigation at time t
$A W H C$	mm	Model input	Available water holding capacity

2.1.2. Thermal time factor

The factor $f_{TT\ t}$ represents the biochemical pathways that are purely controlled by temperature:

$$f_{TT\ t} = \left(\frac{1}{1 + e^{-r_{TT} \left(\frac{T avg_t}{m_{TT}} - 1 \right)}} \right) \cdot 20 \quad (2)$$

where,

$r_{TT} = 5$ is a fixed constant representing the rate that determines steepness of the sigmoid curve.

$m_{TT} = 10$ is a fixed constant representing the midpoint in the sigmoid curve.

$T avg_t$ is the input variable representing observed daily average air temperature at time t for environment k .

In general, crop growth and development rates are expected to increase with temperature. However, once the temperature reaches a point at which the change in conformational states (e.g. denaturation) of enzymes involved in growth and development offsets the increase

in the rates per se, the net effect is a leveling-off of the overall growth and development. Further increases in temperature beyond this inflection point can in some cases result in a decline of the overall growth and development rates, however for the range of temperatures considered here we assume a sigmoidal response for $f_{TT\ t}$ (Eq. (2)). This is consistent with the approach taken by other models such as CropSyst (Stöckle et al., 2003), WOFOST (Van Diepen et al., 1989), and DSSAT-CROPSIM-CERES-Wheat (Jones et al., 2003). These other models make use of piecewise linear functions, however, such functions are not continuously differentiable, a requirement for estimation with Hamiltonian Monte Carlo (HMC). Further explanation of the utility of HMC in the context of crop models is provided below in Section 2.5 and in Poudel et al. (2022). The curve given in Eq. (2) is specified such that the thermal time accumulation is less than 0.2 for temperatures less than zero, has a value of 10 °C-days at 10 °C and approaches a value of 20 °C-days at temperatures at or above 20 °C (Fig. S1 of the Supplementary Materials).

2.1.3. Vernalization factor

For winter wheat, a period of exposure to cold temperatures is required to trigger the transition from vegetative to reproductive growth (Robertson et al., 1996; Ritchie, 1991). This threshold, frequently referred to as the vernalization requirement, varies between cultivars and is represented here by the parameter $vreq$. If the period of exposure to cold temperatures is less than the $vreq$, the transition to reproductive growth (including flowering) will be delayed. To capture this phenomenon we specified a vernalization factor ($f_{vnr\ t}$) calculated as a function of cumulative vernalization at time t (VNR_t ; i.e., the cumulative exposure of the plant to vernalizing temperatures from $t = 0$ to time t), the parameter for vernalization requirement ($vreq$), cumulative thermal time at time t (TT_t) and the parameter TTL :

$$f_{vnr\ t} = \frac{1}{1 + e^{-r_{vnr} \left(\frac{TT_t}{TTS} - 1 \right)}} + \left(1 - \frac{1}{1 + e^{-r_{vnr} \left(\frac{TT_t}{TTS} - 1 \right)}} \right) \left(\frac{1}{1 + e^{-r_{vnr} \left(VNR_t - \frac{vreq}{2} \right)}} \right) \quad (3)$$

where,

TT_t is state variable for thermal time at time t for environment k .

TTS is thermal time to terminal spikelet and assumed to be $\frac{TTL}{2}$ °C-d.

TTL is the parameter representing cumulative thermal time until the end of leaf expansion (i.e. end of vegetative growth), °C d.

$vreq$ is the parameter representing vernalization requirement, vernalization days.

VNR_t is the state variable for cumulative vernalization at time t , calculated below.

r_{vnr} is a constant (assumed to be 100) that specifies the rate of transition from cumulative vernalization having an effect on plant development to having no effect.

r_{vreq} is the rate that determines the shape of the relationship between cumulative vernalization and $f_{vnr\ t}$ and is calculated as $\frac{2 \log(99)}{vreq}$

The equation for $f_{vnr\ t}$ was intended to provide an approximately linear transition from nearly 0 to nearly 1 as VNR_t increased from 0 to $vreq$ (Fig. S2 of the Supplementary Materials). The equation also included a switch function to force $f_{vnr\ t}$ to be approximately 1 when plant development progressed past the terminal spikelet stage (represented here by $TT_t \geq TTS$) based on the fact that the effect of vernalization on the rate of development is only active until that stage (Slafer and Rawson, 1994). The use of sigmoid instead of piecewise linear forms for these relationships was for compatibility with HMC (see discussion in Section 2.5).

2.1.4. Cumulative vernalization

The rate of change of cumulative vernalization (i.e. vernalization rate) was calculated as a sigmoidal function of daily average temperature with vernalization rate at nearly one when temperature was at or below 4 °C and nearly zero when temperature was above 12 °C:

$$\frac{dVNR_t}{dt} = 1 - \frac{1}{1 + e^{-r_{VRN} \left(\frac{Tavg_t}{m_{VRN}} - 1 \right)}} \quad (4)$$

where,

$Tavg_t$ is as defined above.

$r_{VRN} = 5$ is a fixed constant representing the rate that determines the steepness of the sigmoid curve.

$m_{VRN} = 8$ is a fixed constant representing the mid point in the sigmoid curve, i.e. the mid point between 4 °C and 12 °C.

For the range of temperatures addressed in this study, this shape approximates that of other modeling and experimental studies (Wang and Zhao, 2013; Porter and Gawith, 1999). An illustration of this curve is shown in S3 of the Supplementary Materials.

2.1.5. Leaf area index

Leaf area index (LAI) determines the amount of light captured by the crop canopy and, thereby, largely determines the potential growth rate of the crop. The rate of change in LAI ($\frac{dLAI_t}{dt}$) was calculated as the difference between relative net leaf growth (G_{LAI_t}) and senescence (S_{LAI_t}):

$$\frac{dLAI_t}{dt} = G_{LAI_t} - S_{LAI_t} \quad (5)$$

$$G_{LAI_t} = f_{TT\ t} \cdot LAI_t \cdot \alpha \cdot (LAIMAX - LAI_t) \cdot \left(1 - \frac{1}{1 + e^{-r_{LAI} \left(\frac{LAI_t}{LAIMAX} - 1 \right)}} \right) \cdot (1 - f_{gf}) \cdot f_{ws} \quad (6)$$

$$S_{LAI_t} = f_{TT\ t} \cdot LAI_t \cdot senrate \cdot f_{gf} \quad (7)$$

where,

$\frac{dLAI_t}{dt}$ is the rate of change of LAI at time t .

$f_{TT\ t}$ is as described in Eq. (2).

LAI_t is leaf area index at time t .

α is the parameter representing relative rate of LAI growth.

f_{ws} is a factor representing the factor for water stress, calculated as shown below in Eq. (9).

r_{LAI} is the rate that determines the abruptness of the switch to restrict leaf area growth once LAI reaches $LAIMAX$ (assumed to be 100).

$LAIMAX$ is the parameter representing maximum LAI.

f_{gf} is the factor to indicate a switch to grain filling, calculated as shown in Eq. (8).

$senrate$ is the parameter representing the relative rate of senescence per unit thermal time.

Rates of leaf growth and senescence both tend increase with temperature, hence $f_{TT\ t}$ is the first term in Eq. (5). Including LAI_t as a term in Eq. (5) is based on the fact that the potential rates of growth and senescence both depend on current LAI. In the case of growth, this is due to the fact that the photosynthate available for leaf growth is largely dependent on light capture by the current canopy. For senescence, the potential amount of leaf area that can be senesced on a given day is limited by the current leaf area.

In the context of this model, G_{LAI_t} was treated as a combined process (gross growth minus senescence) until the transition to reproductive growth (defined by f_{gf}) after which senescence was assumed to become the dominant process affecting the change in LAI. The use of f_{gf} here forced the effect of $senrate$ on $\frac{dLAI_t}{dt}$ to effectively zero prior to TTL and the effect of G_{LAI_t} on $\frac{dLAI_t}{dt}$ to effectively zero after TT_t exceeded $TTL + 400$ °C-d (meant to represent the beginning of grain filling; see discussion of Eq. (8) below) with a gradual shift from one process to the other in between.

Apart from temperature and current LAI effects, the net leaf growth rate was primarily driven by the parameter for the relative rate of leaf area expansion (α) in units of $(^{\circ}\text{C}\cdot\text{d})^{-1}$. This primary rate was then limited by light interception and soil water availability. At full canopy closure nearly all light will be intercepted at the top of the canopy and older leaves lower in the canopy will senesce due to insufficient light to sustain maintenance respiration (i.e. light intensity falls below the light compensation point). In effect, this imposes a constraint on the maximum LAI that can be sustained by the plant. To represent this maximum LAI constraint two terms were added to Eq. (6): the difference between current LAI and maximum LAI (LAI_{MAX}) and a sigmoidal switch which triggered as LAI_t approached or slightly exceeded the parameter LAI_{MAX} . Finally, because leaf area growth is restricted in the presence of water stress, a water stress factor (f_{ws}) was introduced to slow down leaf growth when soil moisture drops below a given threshold (see Eq. (9) below for the exact calculation).

Figures S5 and S6 in the Supplementary Materials illustrate the effect of current LAI on $\frac{dLAI_t}{dt}$ before anthesis (when G_{LAI_t} dominated) and after anthesis (when $senrate$ dominated).

2.1.6. Grain filling factor

The grain filling factor (f_{gf}) was a phenology-related factor specified to regulate shifts from vegetative to reproductive growth within the model. The equation was specified as a function of TT_t and the parameter representing thermal time to end of leaf expansion (TTL):

$$f_{gf} = \frac{1}{1 + e^{-r_{gf}(TT_t - (TTL + 200))}} \quad (8)$$

where,

TT_t and TTL are defined above.

$r_{gf} = \frac{\log(99)}{200}$ is the rate that controls the shape of the sigmoid curve.

The value for r_{gf} was specified such that Eq. (8) would approximate a piecewise linear function that transitions from 0 at or below $TTL + 100$ $^{\circ}\text{C}\cdot\text{d}$ to 1 at $TTL + 300$ $^{\circ}\text{C}\cdot\text{d}$. By assuming that TTL represents a developmental stage shortly after booting, $TTL + 100$ $^{\circ}\text{C}\cdot\text{d}$ was assumed to be equivalent to 50% heading, $TTL + 200$ $^{\circ}\text{C}\cdot\text{d}$ equivalent to 50% anthesis and $TTL + 300$ $^{\circ}\text{C}\cdot\text{d}$ equivalent to end of grain set (McMaster et al., 2019). An illustration of the curve is provided in S4 of the Supplementary Materials.

2.1.7. Water stress factor

The water stress factor (f_{ws}) was calculated to introduce the effect of water stress on LAI and biomass:

$$f_{ws} = \frac{1}{1 + e^{r_{ws}(\frac{PAW_t}{AWHC} - wsen)}} \quad (9)$$

where,

$wsen$ is a parameter representing water stress sensitivity.

PAW_t is the plant available water at time t .

$AWHC$ is the available water holding capacity.

r_{ws} is the rate that determines the shape of the sigmoid curve and specified equal to $\frac{-\log(99)}{wsen}$ to ensure that f_{ws} was equal to 0.01 at $PAW_t = 0$ and 0.99 at $PAW_t = wsen$.

The value of f_{ws} varied between 0 and 1 following a sigmoidal relationship with soil moisture, where 0 indicated severe water stress and 1 indicated no water stress. The shape of the curve was determined by the parameter $wsen$ which is the inflection point at which f_{ws} is equal to 0.5. Thus, the crop began to experience water stress as soil moisture dropped below $2wsen$. The theory behind this formulation is that for each given soil profile there is a threshold of soil moisture above which water is freely available to the crop. However, as soil moisture drops below this threshold the resistance of soil to the transport of water to the root surface may become limiting resulting in the root system being unable to supply atmospheric demand for transpiration. In such a case, the plant will experience a water deficit, which will inhibit transpiration, leaf area expansion, and plant growth overall.

S7 of the Supplementary Materials shows the curve for f_{ws} with a value of 0.175 for $wsen$ and 150 mm for $AWHC$. For instance, the value of 0.175 for $wsen$ indicates that the crop starts experiencing water stress when at 35% ($2 \times 0.175 = 0.35$) of $AWHC$.

2.1.8. Plant available water content

Plant available water content (PAW) is important for crop growth. Under exposure to inadequate PAW stomata close and thereby shut down crop evapotranspiration, photosynthesis and growth. To represent the water balance for PAW , we propose a single layer lumped model which is relatively simple in terms of model formulation and implementation. The rate of change in PAW_t was calculated as the difference between infiltration and crop evapotranspiration. Infiltration was calculated as a function of precipitation, irrigation, and runoff. Crop evapotranspiration was adjusted to account for water stress.

$$\frac{dPAW_t}{dt} = P_t + I_t - R_t - ET_{c,t} \cdot f_{ws} \quad (10)$$

where,

$\frac{dPAW_t}{dt}$ is the rate of change of plant available soil water content.

P_t is the input variable representing observed precipitation.

I_t is the input variable representing irrigation.

R_t is runoff at time t calculated following the runoff curve number method. The runoff calculations are shown in Eq. S1 to S9 in the Supplementary Materials and follow the procedures in United States Department of Agriculture, Soil Conservation Service, Engineering Division (1986) and Williams (1991).

PAW_t , and f_{ws} are as defined above.

$ET_{c,t}$ is the crop evapotranspiration at time t , calculated as shown below.

2.1.9. Crop evapotranspiration (ET_c)

$$ET_{c,t} = ET_{0,t} \cdot K_{c,t} \quad (11)$$

where,

$ET_{0,t}$ is reference evapotranspiration at time t , mm day^{-1} , calculated following the procedures in Allen et al. (1998) and shown in Eq. S10 to S22 in the Supplementary Materials.

$K_{c,t}$ is the crop coefficient at time t , calculated below.

2.1.10. Crop coefficient (K_c)

The crop coefficient curve described here was constructed to mimic the piecewise linear approach described in Allen et al. (1998):

$$K_{c,t} = K_{c,min,t} + (1 - e^{-K_{LAI,t}}) (K_{c,mid} - K_{c,min,t}) \quad (12)$$

$$K_{c,min,t} = K_{c,ini} - \frac{K_{c,ini} - K_{c,end}}{e^{-r_{Kc}(\frac{TT_t}{TTA} - 1)} + 1} \quad (13)$$

$K_{c,ini}$, $K_{c,mid}$, and $K_{c,end}$ are the crop coefficients at initial, mid-season, and end-season stages and specified as parameters estimated by the model.

TTA is thermal time to anthesis and assumed to be TTL plus 200 $^{\circ}\text{C}\cdot\text{d}$.

r_{Kc} is the rate that determines the shape of the K_c curve (assumed to be 100 to create an abrupt transition)

In Eq. (12), fractional light interception (as calculated based on LAI and K previously shown in Eq. (14)) is used as a proxy for canopy development/decline with low interception representative of $K_{c,t}$ values near the minimum value ($K_{c,min,t}$) and full interception (near 1) representative of the mid-season (maximum K_c) stage. The $K_{c,min,t}$ term itself varied across the season with rapid transition from $K_{c,ini}$ when $TT_t < TTA$ to $K_{c,end}$ when $TT_t > TTA$. The overall effect was that the increase in light interception as a result of leaf area growth early in the season resulted in the transition from $K_{c,ini}$ to $K_{c,mid}$ during what (Allen et al., 1998) termed the crop development stage, and the subsequent decline in interception (due to leaf senescence) resulted in the transition from $K_{c,mid}$ to $K_{c,end}$ during what (Allen et al., 1998) termed the late season stage. Illustration of K_c as a function of LAI before and after anthesis is provided in S9 and S10 of the Supplementary Materials.

2.1.11. Biomass

Biomass at a given point in time is the cumulative prior growth of the crop across the growing season. In this model the rate of change in biomass (i.e. crop growth rate) was calculated as a function of potential RUE, incident solar radiation (SRAD), fractional light interception, temperature and water stress according to the following equation:

$$\frac{dBM_t}{dt} = RUE_p \cdot SRAD_t \cdot (1 - e^{-K \cdot LAI_t}) \cdot f_{TT_t} \cdot f_{ws} \quad (14)$$

where,

$\frac{dBM_t}{dt}$ is the rate of change of biomass at time t .

f_{ws} , f_{TT_t} , and LAI_t are defined above.

K is the parameter representing light extinction coefficient.

RUE_p is the parameter representing potential radiation use efficiency in units of $g\ MJ^{-1}\ (^{\circ}C\text{-}d)^{-1}$

$SRAD_t$ is observed solar radiation at time t .

The amount of growth at a given time is limited by the amount of light intercepted by the crop canopy, which is largely determined by the LAI. Thus, we include the term $1 - e^{-K \cdot LAI_t}$ in Eq. (14) (i.e. the Beer-Lambert equation; [Monsi and Saeki, 2005](#)) to limit growth to the amount of light intercepted by the canopy. The RUE_p parameter represents the combined efficiency of crop photosynthesis and growth and maintenance respiration with which a given crop can produce biomass per unit of intercepted solar radiation per $^{\circ}C\text{-}d$. The product of RUE_p , $SRAD_t$, and fractional light interception can be considered the potential growth rate $(^{\circ}C\text{-}d)^{-1}$. By further multiplying this value by f_{TT_t} we achieved the temperature-adjusted biomass growth rate. The multiplication by f_{ws} imposed a further soil moisture constraint on biomass growth following the reasoning that if water stress is sufficient to reduce transpiration then photosynthesis (and subsequent growth) would also be correspondingly reduced. S8 of the Supplementary Materials shows how biomass changes with respect to LAI when everything else is held constant.

2.1.12. Grain weight

The grain growth rate is the dominant form of growth during late reproductive growth (i.e. grain-filling period). The majority of biomass increase during this period is allocated to the grain. As such, the rate of change of grain weight was calculated as the product of biomass growth rate at time t and f_{gf} :

$$\frac{dYLD_t}{dt} = \frac{dBM_t}{dt} \cdot f_{gf} \quad (15)$$

where,

$\frac{dYLD_t}{dt}$ is the rate of change of grain yield at time t .

$\frac{dBM_t}{dt}$ and f_{gf} are defined above.

The assumption of this approach is that all grain weight is directly derived from photosynthesis during the grain-filling period. While this does not represent the full biological reality, it was intended as a first-order approximation suitable for the present application.

2.1.13. Full ODE model

[Fig. 1](#) provides a Forrester diagram that illustrates the full ODE model. From this diagram it is clear that temperature is a major driver of the model. The f_{TT_t} factor, which is calculated from T_{avg_t} , directly or indirectly affects terms in the rate of change of four of the six state variables (TT_t , LAI_t , BM_t and YLD_t). VRN_t also depends on temperature and affects the rate of change of TT_t through f_{vrn_t} , but both of these state variables are otherwise independent of other parts of the model. TT_t affects the rest of the model through f_{gf} , which marks the transition from vegetative to reproductive growth and thereby ceases leaf growth (G_{LAI_t}), initiates leaf senescence (S_{LAI_t}) and allocation of biomass toward grain yield ($\frac{dYLD_t}{dt}$). Temperature also has an indirect effect on PAW_t through the effect of TT_t on ET_{c_t} . The effect of water stress is transferred through the f_{ws} factor to G_{LAI_t} , which creates a balancing loop in which f_{ws} reduces G_{LAI_t} , which restricts LAI_t , which then restricts ET_{c_t} . This f_{ws} factor directly affects $\frac{dBM_t}{dt}$,

but also has an indirect effect on $\frac{dBM_t}{dt}$ by restricting light interception through the effect of f_{ws} on LAI_t . Both of these effects on $\frac{dBM_t}{dt}$ have potential to affect $\frac{dYLD_t}{dt}$ if they occur during the reproductive phase of growth.

2.2. Euler integration

At each time step, the value of the state variables were updated for each k th environment using Euler integration according to:

$$S_{k\ t+1} = S_{k\ t} + \frac{dS_{k\ t}}{dt} \Delta t \quad (16)$$

where, $S_{k\ t} = (VRN_{k\ t}, TT_{k\ t}, LAI_{k\ t}, BM_{k\ t}, YLD_{k\ t}, PAW_{k\ t})$ are the observed state variables for environment k at time t , $S_{k\ t+1} = (VRN_{k\ t+1}, TT_{k\ t+1}, LAI_{k\ t+1}, BM_{k\ t+1}, YLD_{k\ t+1}, PAW_{k\ t+1})$ at time $t+1$, and Δt is the fixed time step of one day.

The output of the ODE model is then represented by:

$$\beta_{0ikt} = f(\mathbf{I}_k, \boldsymbol{\phi}, S_0) \quad (17)$$

where, f represents the numerical integration of the ODE system from time 0 to time t , $\mathbf{I}_k$ is a matrix of input variables, $\boldsymbol{\phi}$ is a vector of the ODE model parameters as described above, and S_0 is the vector of state variables at time 0.

2.3. Integration of the ODE model into a Bayesian hierarchical framework

The ODE model was integrated into a Bayesian hierarchical framework as follows for LAI, biomass, and PAW:

$$Y_{ijkt} = \beta_{0ikt} + Plot[Env]_{j[k]} + \epsilon \quad (18)$$

where,

Y_{ijkt} is the observed value of the i th response variable ($i = 1$ for LAI, $i = 2$ for biomass, and $i = 3$ for PAW), corresponding to j th plot in the k th environment at time t . To recall, an environment is defined as a location-year-treatment combination.

β_{0ikt} is the predicted value of the i th response variable from the ODE model for environment k and time t .

$Plot[Env]_{j[k]}$ is the differential effect of j th plot within an environment on the i th response variable, assumed $Plot[Env]_{j[k]} \sim NIID(0, \sigma_{Plot\ i}^2)$.

ϵ is residual error, assumed $\epsilon \sim NIID(\sigma_{\epsilon}^2)$.

For grain yield, the plot effect was excluded as it contains information at only the final time point. Thus, the equation for grain yield was:

$$Y_{ijkt} = \beta_{0ikt} + \epsilon \quad (19)$$

where, Y_{ijkt} , ϵ are as defined above with $i = 4$ for grain yield, and β_{0ikt} is the state variable for grain weight ($i = 4$) for environment k at $t =$ time of harvest.

The Bayesian approach allowed us to quantify uncertainty around the parameters for both the dynamic model and the hierarchical statistical model. It also allowed us to incorporate pre-existing knowledge about the parameters into the model in the form of prior specification, which is especially useful in the case of dynamic model parameters because of their biological meaning.

2.4. Prior specification

The prior distributions for the ODE model parameters except for $senrate$ were estimated based on literature and were specified as truncated normals in the form of $N(\mu, \sigma^2)$. These parameters were specified to have a lower bound of zero and no upper bound.

$\alpha \sim N(0.016, 0.007^2)$ truncated at zero ([Rodriguez et al., 1998](#)).

$K \sim N(0.6, 0.1^2)$ truncated at zero ([Bechini et al., 2006](#); [Muurinen and Peltonen-Sainio, 2006](#)).

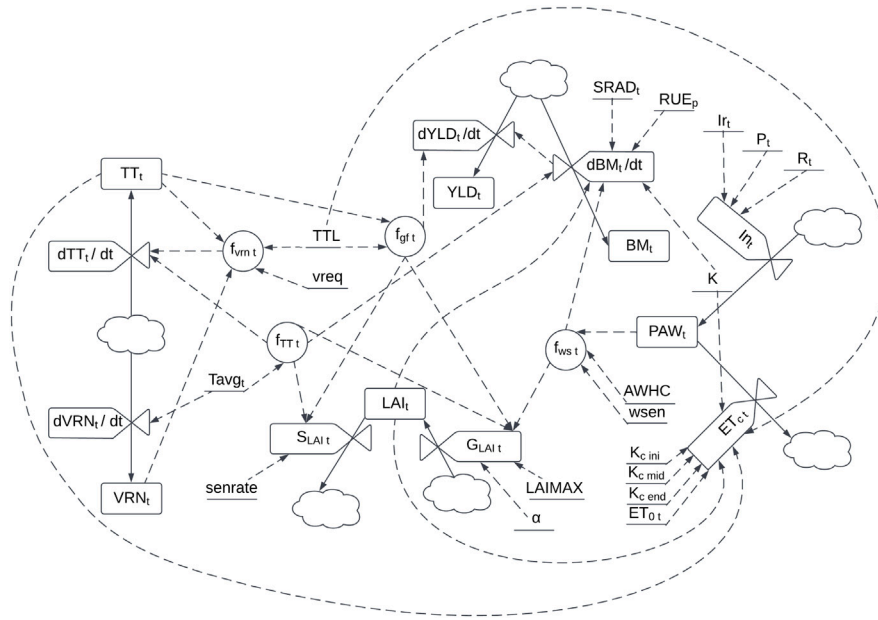


Fig. 1. Forrester diagram of the full ordinary differential equations model including state variables for thermal time (TT_t), cumulative vernalization (VRN_t), leaf area index (LAI_t), crop biomass (BM_t), grain yield (YLD_t) and plant available water content (PAW_t). State variables are represented as rectangles. Parameters and input data are underlined. Flow of material is represented by solid arrows. Flow of information is indicated by dashed arrows. Terms related to the rate of change of state variables are represented by the valve shape. Time-dependent intermediate factors that are calculated from a combination of state variables, parameters and input data are represented as circles. Further detail on the definition of parameters, input data and equations are provided in Table 1 and Sections 2.1.1 to 2.1.12.

$TTL \sim N(950, 100^2)$ truncated at zero (McMaster et al., 2019; Bechini et al., 2006).

$RUE_p \sim N(0.07, 0.0125^2)$ truncated at zero (Bechini et al., 2006; Muirinen and Peltonen-Sainio, 2006). The units for RUE_p were changed from $g MJ^{-1}$ reported in literature to $g MJ^{-1} ^\circ C^{-1}$ assuming an optimal temperature of $20 ^\circ C$.

$vreq \sim N(42, 7^2)$ truncated at zero (Li et al., 2013; Crofts, 1989).

$LAIMAX \sim N(5, 2^2)$ truncated at zero (Wagle et al., 2021; Nielsen et al., 2012).

The prior for $senrate$ was determined using a heuristic procedure. It was assumed that with a peak LAI at flowering of 5 and $30 ^\circ C$ daily average temperature, it would take approximately 35 days to reach full senescence. Values of $senrate$ were tested heuristically to determine a range of values that resulted in plausible senescence durations. With a rate of senescence of 0.005, it took approximately 35 days to reach an LAI near zero from five, a $senrate$ of 0.002 took approximately 60 days, and a $senrate$ of 0.008 took approximately 20 days. Thus, the prior μ was specified as 0.005 with a sigma of 0.001, thereby providing support in the prior for a range of duration between 20 and 60 days within three standard deviations of prior μ :

$senrate \sim N(0.005, 0.001^2)$ truncated at zero.

The prior means for the parameters pertaining to the water balance component ($K_{c\ ini}$, $K_{c\ mid}$, $K_{c\ end}$, and w_{sen}) were also specified based on literature. The standard deviations were chosen such that the prior allows for all practically possible parameter values without being too vague. The priors were specified as:

$K_{c\ ini} \sim N(0.7, 0.2^2)$ truncated at zero (Allen et al., 1998).

$K_{c\ mid} \sim N(1.15, 0.1^2)$ truncated at $K_{c\ ini}$ (Allen et al., 1998).

$K_{c\ end} \sim N(0.25, 0.05^2)$ truncated at zero with upper bound $K_{c\ mid}$ (Allen et al., 1998).

$w_{sen} \sim N(0.175, 0.025^2)$ truncated at zero with upper bound 1 (Amir and Sinclair, 1991).

The priors for the hyperparameters pertaining to the hierarchical component of the model were specified following the prior predictive procedure described in Schad et al. (2019):

$\sqrt{\sigma_{Plot\ LAI}^2} \sim N(0, 0.1^2)$ truncated at zero.

$\sqrt{\sigma_{Plot\ biomass}^2} \sim N(0, 0.1^2)$ truncated at zero.

$\sqrt{\sigma_{Plot\ PAW}^2} \sim N(0, 0.1^2)$ truncated at zero.

$\sqrt{\sigma_e^2 LAI} \sim N(0, 1^2)$ truncated at zero.

$\sqrt{\sigma_e^2 biomass} \sim N(0, 100^2)$ truncated at zero.

$\sqrt{\sigma_e^2 PAW} \sim N(0, 50^2)$ truncated at zero.

$\sqrt{\sigma_e^2 yield} \sim N(0, 30^2)$ truncated at zero.

2.5. Model implementation and sampling

Two separate versions of the model were implemented, one with water balance (+WB) and another without water balance (-WB) to assess the contribution of the water balance component on the predictive ability of the model. The +WB model included Eq. (10) and (11), whereas these equations were excluded in the -WB model, and f_{ws} was fixed at 1 indicating no water stress.

The data used for parameter estimation were obtained from a previous study by Lollato and Edwards (2015). It includes observed data at multiple time points throughout a growing season for LAI, biomass, PAW, and yield from multiple environments. Weather data for average daily temperature, precipitation, solar radiation, wind speed at 2 m, relative humidity, and atmospheric pressure were obtained from Oklahoma Mesonet stations (McPherson et al., 2007). A detailed description of the data is presented in the supplementary materials and Lollato and Edwards (2015).

Samples were drawn from the posterior distribution of interest using a dynamic HMC sampler as implemented within Stan (Stan Development Team, 2020b). Dynamic HMC is a type of Markov Chain Monte Carlo (MCMC) algorithm that utilizes gradient-based sampling of the typical set of the joint posterior density (Betancourt, 2017). This gradient-based approach has been shown to efficiently sample high-dimensional parameter spaces with complex geometries (Betancourt and Girolami, 2015) similar to the complex interactions between the parameters of a dynamic crop simulation model. Because HMC utilizes the gradient for sampling, all model components must be continuously differentiable so that the gradient can be calculated at any point in the parameter space. Although this imposes a constraint on model forms

that can be used, the gains in sampling efficiency were considered to be worth the inconvenience of adopting HMC-compatible functional forms. Further discussion regarding HMC and its application in crop modeling can be found in Poudel et al. (2022).

The models were implemented in Stan version 2.25.0 (Stan Development Team, 2020b) with the command-line interface to the Stan modeling language, CmdStan version 2.25.0 (Stan Development Team, 2020a). Four MCMC chains were run with 20,000 iterations including 50% burn-in, resulting in a total of 40,000 saved iterations. Sampling was performed at the Oklahoma State University High Performance Computing Center (OSU-HPCC) using the Pete supercomputer and post-sampling analysis was performed on a Linux virtual machine hosted on The Interactive Graphical Environment for Research (TIGER) research cloud at OSU-HPCC. Two common convergence diagnostics, traceplots and R-hat values, were used to monitor chain convergence. The effective sample size (ESS) for the ODE model parameters and the hierarchical model hyperparameters were greater than 6,000 in both models.

For cleaning and organizing the data, R package tidyverse was used (Wickham et al., 2019; Wickham, 2017). The posterior samples were processed with the R statistical software environment (R Core Team, 2020). The highest density intervals (HDIs) of the posterior distributions of the parameters were computed using the HDInterval package (Meredith and Kruschke, 2018). The tables were generated using the R packages knitr (Xie, 2020) and kableExtra (Zhu, 2019). The figures were created using the ggplot2 (Wickham, 2016), gridExtra (Auguie, 2017) packages in R.

2.6. Model predictive performance

A cross-validated dataset was generated with leave-one-group out cross-validation where each environment was considered a group, hence there were a total of ten groups. For each iteration of the cross-validation, one group was held out and parameters were estimated based on the remaining nine groups. The parameter estimates were then used to predict the data from the withheld group. The predictions were used to calculate three statistical metrics:

2.6.1. Relative Root Mean Square Error ($rRMSE$):

$$rRMSE = \frac{RMSE_s}{\bar{y}} \quad (20)$$

$$RMSE_s = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n^s)^2} \quad (21)$$

where, N = Total number of data points, y_n is the n th observation ($n=1,2,\dots,N$), $\bar{y}$ is the average of the observed data points, and $\hat{y}_n^s$ is the predicted value for the n th observation obtained from the s th MCMC iteration. Models with smaller values of RMSE are preferable.

2.6.2. Willmott agreement index (d ; Willmott, 1981):

$$d_s = 1 - \frac{\sum_{n=1}^N (y_n - \hat{y}_n^s)^2}{\sum_{n=1}^N (|\hat{y}_n^s - \bar{y}| + |y_n - \bar{y}|)^2} \quad (22)$$

where, N , y_n , and $\hat{y}_n^s$, and $\bar{y}$ are as described above. This statistic ranges between 0 to 1 with values closer to 1 indicating good model fit.

2.6.3. Nash–Sutcliffe Efficiency (NSE ; Nash and Sutcliffe, 1970):

$$NSE_s = 1 - \frac{\sum_{n=1}^N (y_n - \hat{y}_n^s)^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \quad (23)$$

where, N , y_n , $\hat{y}_n^s$, and $\bar{y}$ are as described above. The values of NSE can range from $-\infty$ to 1 and values closer to 1 indicate a better-fitting model.

This process was repeated for all groups one at a time and for both versions of the model i.e. with and without the water balance component. Model comparison was done to identify how water balance influences estimations and predictions of different state variables.

3. Results

3.1. Model comparison for LAI, biomass, PAW, and yield predictions

To quantify the differences between the two models, Fig. 2 shows the density plots of the prediction statistics, $rRMSE$, d , and NSE computed under the 10-fold cross validation. Fig. 2 shows that adding the water balance component had a larger impact on grain yield compared to LAI or biomass as indicated by differences in the prediction statistics between the models. The posterior median values of $rRMSE$, d , and NSE for the +WB model were 0.67, 0.89, and 0.62 for LAI, 0.32, 0.98, and 0.92 for biomass, 0.52, 0.75, and -0.03 for PAW, and 0.23, 0.9, and 0.4 for grain yield. The +WB model version had higher values for NSE and d and lower values for $rRMSE$ for all variables. For example, median NSE increased from 0.57 to 0.62 for LAI, 0.74 to 0.92 for biomass, and -0.58 to 0.40 for grain yield.

Fig. 3 shows median values for cross-validation predicted grain yield from both models along with observed yield data. The model predicted grain yields for rainfed environments differed substantially from observed values. The over-prediction of yield was most evident in cases of rainfed environments, particularly in 2014 (Fig. 3) when the crop experienced severe water deficit stress. Unlike biomass, the difference between estimated and predicted yields were negligible for yield. Furthermore, the yield values predicted by the -WB model were clustered between 500 to 600 g m⁻² for all environments (Fig. 3). Contrastingly, median yields predicted by the +WB model ranged between 100 and 800 g m⁻² depending on environment. Taken together, the resulting trend across environments was qualitatively more similar to the observed data (i.e. correspondence between low versus high yielding environments) for the +WB model than the -WB model.

3.2. Joint posterior distribution of ODE model parameters

Fig. 4 shows the correlation between the ODE model parameters from both models. Varying degrees of correlations were observed between the posterior samples of the ODE model parameters. Specifically, correlations were observed between α and LAI_{MAX} ($\hat{r} = -0.96$), TTL and $vreq$ ($\hat{r} = -0.8$), $RU E_p$ and K ($\hat{r} = -0.66$), and $senrate$ and K ($\hat{r} = 0.52$) in the +WB model. Similar correlation pattern followed among these parameters in the -WB model. In addition, correlations were observed between $RU E_p$ and $K_{c mid}$ ($\hat{r} = 0.83$), K and $K_{c mid}$ ($\hat{r} = 0.65$), $senrate$ and $K_{c mid}$ ($\hat{r} = 0.53$), $K_{c mid}$ and $wsen$ ($\hat{r} = 0.50$), and $RU E_p$ and $wsen$ ($\hat{r} = 0.51$) in the +WB model. The biggest differences in the posterior densities of the two models were observed in $RU E_p$, K and $senrate$. The parameters $RU E_p$ and K were estimated to be higher whereas $senrate$ was lower for the model including water balance compared to the -WB model.

3.3. Crop growth parameters from the ODE component of the model

Table 2 shows posterior HDI and median for the crop growth parameters from the +WB model. The HDI for α ranged from 0.0010 to 0.0012 with a median of 0.0011, which indicates that the rate of LAI increase is 0.11% per °C-d before TTL . For a 20 °C-d, that would amount to 2.2% of LAI increase per day. Likewise, the posterior median for the canopy light extinction coefficient (K) was estimated to be 0.46. The model estimated that it takes around 737 °C-d to reach end of leaf expansion as indicated by the median for TTL , and the average vernalization requirement as indicated by the posterior median for $vreq$ was 42.8 days. The parameter $senrate$ had a median of 0.0053, which indicates that it takes 36 days for the canopy to reach from maximum

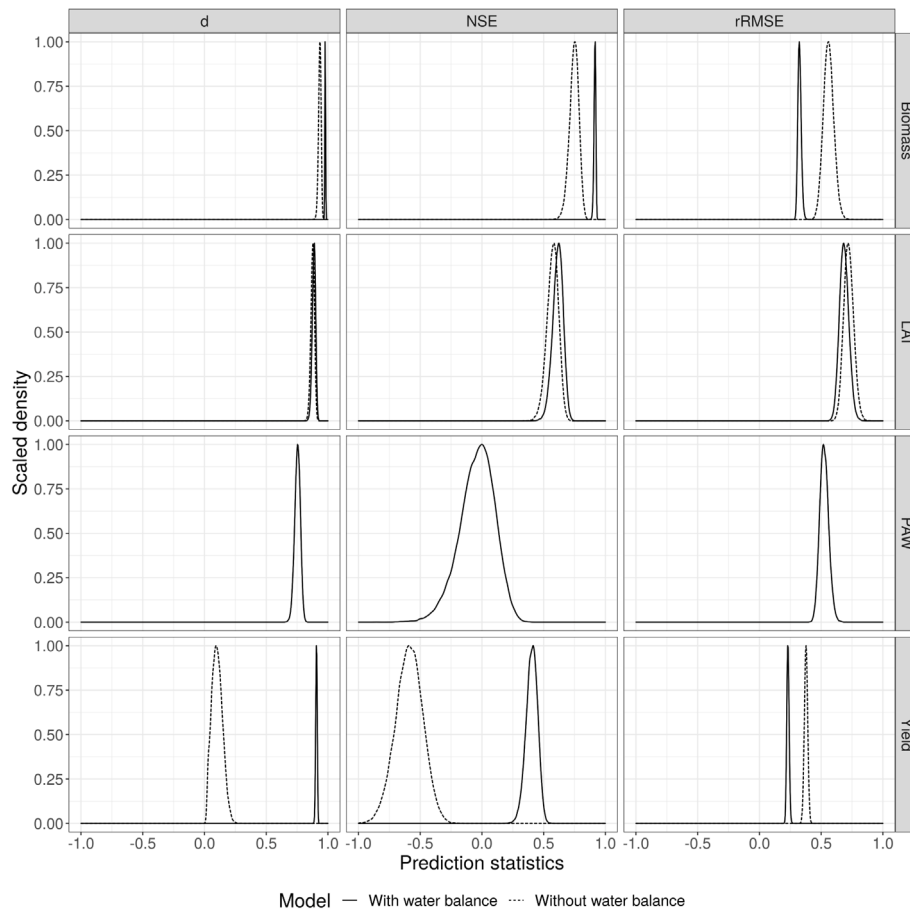


Fig. 2. Predictive performance of models including and excluding the water balance component for leaf area index (LAI), biomass, plant available water (PAW), and grain yield assessed with three statistical metrics: relative root mean squared error (rRMSE), Willmott agreement index (d), and Nash–Sutcliffe efficiency (NSE).

Table 2

Posterior HDI and median for the ODE model parameters related to crop growth obtained from the model with water balance.

Parameter	Description	HDI	Median
α	Relative rate of LAI increase before the end of leaf expansion	(0.0010, 0.0012)	0.0011
K	Light interception coefficient	(0.38, 0.55)	0.46
TTL	Thermal time to the end of leaf expansion	(693, 784)	737
$senrate$	Rate of senescence	(0.0049, 0.0057)	0.0053
RUE_p	Potential radiation use efficiency	(0.100, 0.117)	0.109
$vreq$	Vernalization requirement	(28.8, 56.7)	42.8
$LAIMAX$	Maximum leaf area index	(6.06, 7.24)	6.66

leaf area to an LAI of less than 0.1 at a daily average temperature of 30 °C. Radiation use efficiency is more commonly expressed in units of $g\ MJ^{-1}$, which could be obtained by multiplying RUE_p by 20 °C-d which would result in potential RUE ($g\ MJ^{-1}$) with median of 2.17 and HDI of (1.99, 2.34).

Fig. 5 shows the smoothed densities of posterior and prior distribution samples of the crop growth parameters from the +WB model. The parameters α , K , and TTL shifted lower than specified prior,

whereas RUE_p shifted higher, which indicates that the data were informative in estimating the parameters. Likewise, the densities of posterior distributions were considerably narrowed compared to that of prior distributions for all parameters except $vreq$ showing that the data helped reduce uncertainties around those parameters.

3.4. Water balance parameters from the ODE component of the model

Fig. 6 shows the smoothed densities of samples from prior and posterior distributions for the water balance parameters. The posterior density curve for $K_{c\ ini}$ shifted much lower than specified prior, whereas the posterior density curves for $K_{c\ mid}$, $K_{c\ end}$, and w_{sen} shifted higher compared to their prior density curves. The posterior HDI and median value for $K_{c\ ini}$, $K_{c\ mid}$, $K_{c\ end}$, and w_{sen} were (0.000004, 0.08) and 0.03, (1.27, 1.47) and 1.37, (0.23, 0.42) and 0.32, and (0.21, 0.25) and 0.26 respectively.

3.5. Parameters from the hierarchical component of the model

As we have established that the +WB model had better goodness-of-fit and predictive ability, this and the following sections will present results from the +WB model for making inferences.

Table 3 shows the posterior HDI and median for the plot effect within an environment ($Plot[Env]_{ij[k]}$) and the residual variance $\sqrt{\sigma_e^2}$ for LAI, biomass, and PAW. The residual variance indicates the amount of variance unexplained by the model. The plot level effects are the

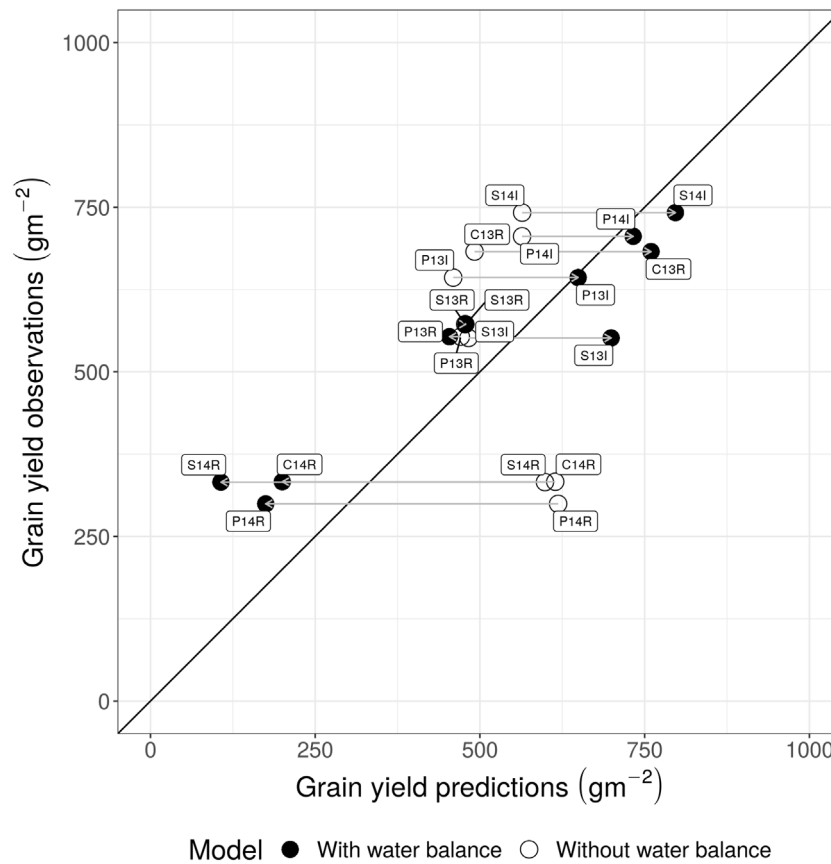


Fig. 3. Cross-validated median estimations and predictions of grain yield comparing models with and without a water balance component to the mean observed data points for winter wheat grown in Oklahoma at Chickasha (C), Stillwater (S), and Perkins (P) across the 2012–2013 (13) and 2013–2014 (14) seasons under irrigated (I) and rainfed (R) conditions. Data points are uniquely identified by a four-digit code LYYM, where L represents the location (C, S, or P), YY is the two-digit year (13 or 14) and M represents management (I or R). Gray arrows show the shift in predictions that result from adding the water balance component.

Table 3

Posterior HDI and median for the parameters from the hierarchical component in the model.

Parameter	HDI	Median
$Plot[Env]_{i j k}$ for LAI	(−0.51, 0.47)	0.010
$Plot[Env]_{i j k}$ for biomass	(−0.14, 0.18)	0.002
$Plot[Env]_{i j k}$ for PAW	(−0.41, 0.41)	−0.150
$\sqrt{\sigma_e^2 LAI}$	(0.85, 0.98)	0.920
$\sqrt{\sigma_e^2 biomass}$	(152, 176)	164.000
$\sqrt{\sigma_e^2 suw}$	(24.86, 28.63)	26.680
$\sqrt{\sigma_e^2 yield}$	(86.90, 125.33)	105.000

differential effects of plots within an environment. As mentioned previously, the plot effects were defined as a fraction of the ODE-model-simulated value for each variable. Thus, the plot effects with median closer to zero indicate an absence of model bias whereas the plot effects deviating from zero indicate a model bias. The posterior median for the plot effects were close to zero for LAI and biomass, whereas we observed a negative bias for PAW with a posterior median of −0.15.

4. Discussion

In this study, we utilized an ODE crop growth model coupled with a simple water balance model to estimate certain wheat growth and water balance parameters in winter wheat.

4.1. Model comparison for LAI, biomass, PAW, and yield predictions

Comparison between the models with (+WB) and without (-WB) the water balance component showed mixed results depending on the variable being considered (Fig. 2). The greatest improvement in model performance (from -WB to +WB) was seen for yield. Unexpectedly, the results for LAI showed very little difference in model performance between the two versions. With the +WB model, biomass showed the best model performance, while PAW ranked worst, when comparing across all three statistical metrics ($rRMSE$, d , NSE). For all variables NSE indicated poorer performance than d , most notably in the case of PAW. The NSE statistic is a comparison against the overall mean of observations and is, thus, especially sensitive to model bias. The d statistic is better tuned to assess the extent to which a model captures the overall dynamics of observed data (e.g. the timing of increasing vs. decreasing trends). These results suggest that the +WB model version is representing the overall dynamics of the system well, even if some overall bias is present.

Despite the overall similar performance for LAI and biomass, the model performance for grain yield was drastically different for +WB and -WB (Figs. 2 and 3). The finding that grain yield predictions were significantly improved by adding soil water balance aligns with previous research indicating the importance of water availability for wheat yield performance in this region (Berhe et al., 2017; Lollato et al., 2017, 2020; Sciarresi et al., 2019). However, the discrepancy between small improvements in LAI and biomass predictions and the substantial improvement in yield is noteworthy.

The grain yield predictions by the -WB model did not capture the general trend in grain yield across environments as shown by the lack of conformity with the one to one line in Fig. 3. Specifically, the -WB

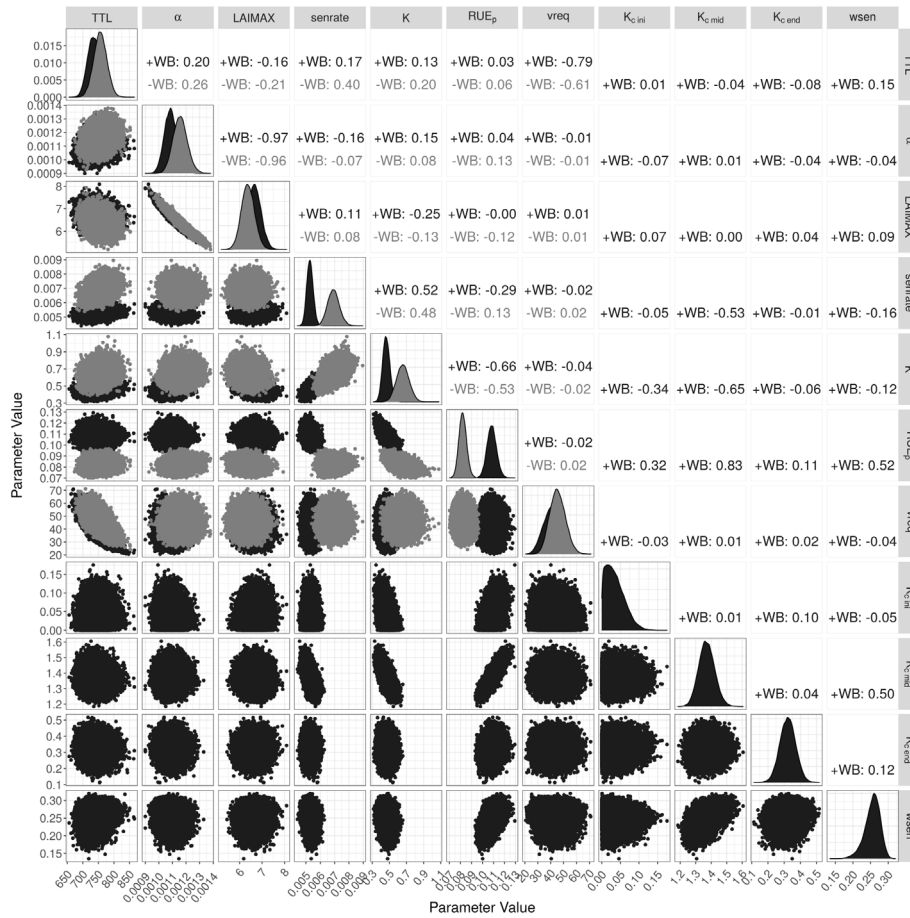


Fig. 4. Joint posterior summary of the ODE model parameters, namely, relative rate of LAI increase before the end of leaf expansion (α), light extinction coefficient (K), Thermal time to end of leaf expansion (TTL), rate of senescence ($senrate$), potential radiation use efficiency (RUE_p), vernalization requirement ($vreq$), maximum LAI ($LAIMAX$), crop coefficient during initial growth stage ($K_{c\ ini}$), crop coefficient during mid-season ($K_{c\ mid}$), crop coefficient during end-season ($K_{c\ end}$), and water stress sensitivity parameter ($wsen$) for the model with water balance (+WB; black) and the model without water balance (-WB; gray). The upper triangle shows correlation coefficients between the parameters for +WB and -WB models.

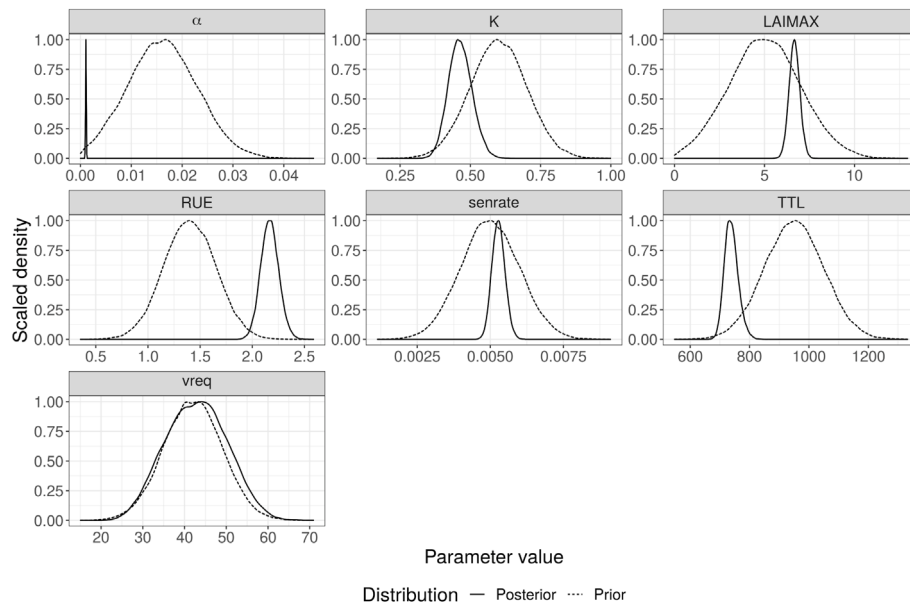


Fig. 5. Density of prior and posterior distribution from the model with water balance for the ODE model parameters related to crop growth namely, relative rate of LAI increase before the end of leaf expansion (α), light extinction coefficient (K), thermal time to end of leaf expansion (TTL), relative rate of senescence ($senrate$), potential radiation use efficiency (RUE_p), and vernalization requirement ($vreq$).

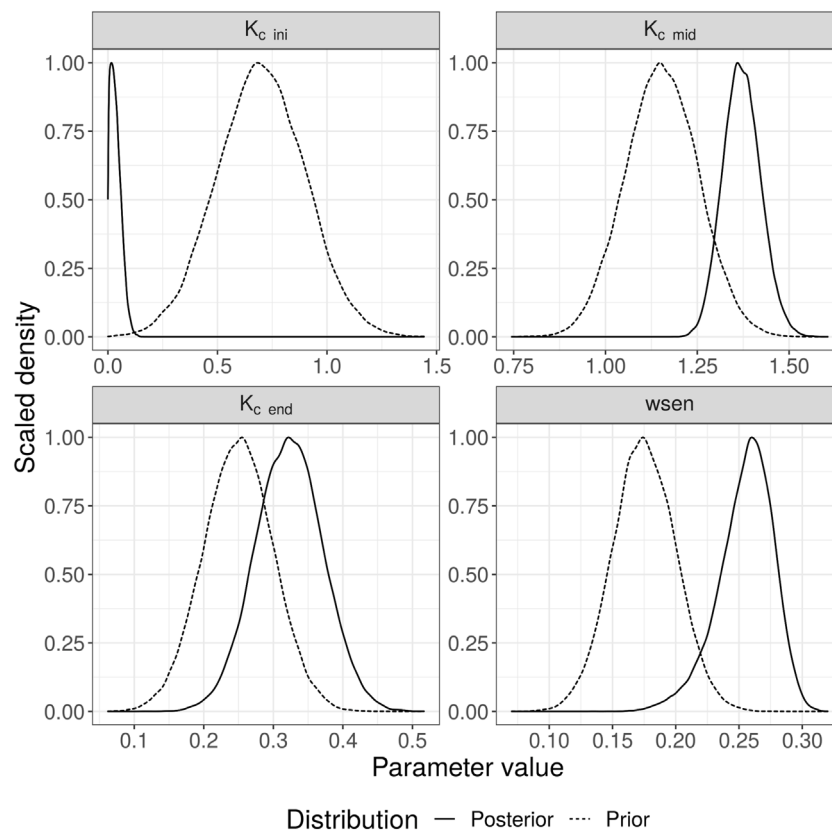


Fig. 6. Density of prior and posterior distribution of the water balance component parameters namely, crop coefficient during initial growth stage ($K_{c\ ini}$), crop coefficient during mid-season ($K_{c\ mid}$), crop coefficient during end-season ($K_{c\ end}$), and water stress sensitivity parameter (w_{sen}).

model substantially over-predicted the grain yield in rainfed environments in 2014 (Fig. 3), when low rainfall resulted in severe water stress late in the season. This severe water stress was captured by the +WB model, as evidenced by the leftward shift of yield predictions when water balance was incorporated into the model (Fig. 3). In contrast, the rightward shift in grain yields for low water stress environments (most sites in the 2013 season and irrigated sites in 2014 season) represented a release of the constraint previously imposed by the -WB model. In other words, the -WB model provided no means to distinguish between high and low water stress environments thereby forcing the sampler to maximize the likelihood by estimating values for model parameters that resulted in yield predictions clustered around the average across all environments. This phenomenon was present even when data from all environments was used for parameter estimation (data not shown).

For the +WB model, the simulated water stress effects restricted growth in water-stressed environments thereby allowing the sampler to estimate values for parameters that resulted in higher yield predictions for environments without water stress (Fig. 3). With the inclusion of water balance and the concomitant shifts in predicted yield (higher for irrigated sites and lower for rainfed sites), the model-predicted trend more closely matched that of the observed data, as indicated by the proximity of the predicted points to the one to one line. Nevertheless, grain yield predictions by the +WB model show some room for improvement, particularly for rainfed sites in 2014 where predictions were lower than observations indicating that the model may have overestimated water stress. Unlike biomass, the grain yield predictions for +WB differed noticeably from observed values (Fig. 3). This difference in pattern between biomass and grain yield shows that the yield estimation was limited despite seemingly accurate biomass estimation.

Another contributing factor might be the fact that the model does not represent drought response mechanisms such as enhanced root growth, stomatal closure or carbohydrate remobilization, any of which

might have come into play to affect the yield formation in the actual crop. In particular, carbohydrate remobilization can play an important role in grain-filling under water-limited conditions (Yang et al., 2000; Davidson and Chevalier, 1992). Thus, a possible improvement in the model to address this issue could be to represent the carbohydrate remobilization process.

4.2. Crop growth parameters from the ODE component of the model

The model yielded biologically plausible values for most parameters, except RUE_p , for which the estimated values were higher than plausible. The posterior median for the light extinction coefficient parameter (K) was estimated to be 0.46 which is biologically plausible but lower than most literature reported values that were used to formulate the priors. A study by Pradhan et al. (2018) in India obtained light extinction coefficient values between 0.47 to 0.65 under irrigated conditions, and reported that the irrigation did not have a significant effect on light extinction coefficient. Another study by Kukal and Irmak (2020) in Nebraska, USA obtained light extinction coefficient values of 0.36 and 0.66 across two growing seasons for wheat. Low light extinction coefficient indicates that it requires more leaf area to intercept the same amount of light compared to canopies with high light extinction coefficient. The parameter K also negatively correlated with RUE_p ($r = -0.66$). The estimated RUE expressed in $g\ MJ^{-1}$ (as calculated in Section 3.3) was substantially higher than the range of values (0.7 to 1.8 $g\ MJ^{-1}$) reported by other studies (Caviglia et al., 2004; Albrizio and Steduto, 2005). Notably, the posterior density for RUE_p also shifted higher than the specified prior distribution. At first glance, it seemed that the water stress was overestimated by the model since the RUE_p from the -WB model was within a biologically plausible range (Fig. 4) and the grain yields for rainfed environments were predicted to be lower than observed for 2014, the year with severe water stress [See

Fig. 3 of the Supplementary Materials]. Given that no such decrease was observed in biomass [See Fig. S11 of the Supplementary Materials], it may indicate that the higher-than-biologically-plausible RUE_p values may have compensated for the downward pressure on biomass due to water stress. However, the temperature effect on biomass growth may also have been a contributing factor. The ODE model in its current form assumes the same temperature function for phenology and biomass growth. However, Wang et al. (2017) reviewed recent research suggesting that the range of optimum temperatures for biomass growth may be wider than that for phenological development or photosynthesis alone. In our case, the higher values estimated for RUE_p may have also been compensating for the negative effect of cool temperatures on simulated growth during the early season. Hence, specifying separate temperature response functions for phenology and biomass growth may produce better parameter estimates for RUE_p .

The correlations between the ODE model parameters indicate some difficulties in uniquely identifying them (Fig. 4). For example, the interplay between TTL and $vreq$ indicates that there is not enough information in the data to adequately distinguish these two parameters. The dataset used in this study does not adequately represent diverse conditions in terms of planting time and temperatures which would be helpful in identifying vernalization requirement. A dataset from environments with variable planting time and more contrasting temperatures could potentially provide information to disambiguate these parameters. For instance, a planting date ranging from September to November would encompass a greater range of planting conditions, which is not outside the bounds of usual planting dates for the study region (Munaro et al., 2020). Likewise, the strong negative correlation between α and $LAIMAX$ ($\hat{r} = -0.96$) indicates that the information provided in the data was not sufficient to disambiguate these parameters. From a biological perspective, this finding is not especially surprising as these two parameters represent closely-related phenomena. Future model development may involve simplifying LAI-related equations to allow estimation of a single leaf growth parameter in place of α and $LAIMAX$.

The parameters also showed considerable differences in the posterior distributions compared to their prior distributions except $vreq$. A heavy shift and/or narrowing down of the density curve from prior to posterior indicates that the data were informative in estimating the parameters. For instance, the time to end of leaf expansion (TTL) was estimated to be shorter (posterior median of 737 °C-d) than our prior expectation. Likewise, the parameter α was estimated to be considerably lower than the prior we specified. The posterior density curve for the parameter $vreq$ mostly overlapped with its prior density curve which indicates that there was essentially no additional information regarding vernalization requirement in the data. This is most likely because the planting dates across these environments were similar and there is a robust prior knowledge to appropriately represent the vernalization requirement with these typical planting dates.

4.3. Water balance parameters from the ODE component of the model

The water balance parameters ($K_{c\ ini}$, $K_{c\ mid}$, $K_{c\ end}$, and $wsen$) were estimated to be different from the priors as indicated by the differences in prior and posterior samples density curves (Fig. 6). The parameter $K_{c\ ini}$ shifted much lower with a median of 0.03 compared to prior mean of 0.7. It indicates that the crop coefficient during early season is much lower than the tabulated value in FAO-56 for Oklahoma conditions. A study by Ko et al. (2009) to determine crop coefficients with the lysimeter method reported that K_c values ranged from 0.1 to 1.7 across the growing seasons in south Texas conditions. By contrast, the estimates for $K_{c\ mid}$ were higher (median of 1.37) than the specified priors pointing towards higher evapotranspiration than expected. The high $K_{c\ mid}$ estimates could partially be due to the functional form of the model whereby the light interception asymptotically approaches 1 but never reaches 1. The parameter $K_{c\ mid}$ was positively correlated

with the crop growth parameters RUE_p and $wsen$ and negatively correlated with K and $senrate$. These correlations suggest that the relationship between water stress and crop growth can be manifested through RUE_p , K , and $senrate$.

The posterior density curve for $K_{c\ end}$ also shifted higher but with a great overlap with the prior density. The water stress sensitivity parameter ($wsen$) for the sigmoid curve to calculate f_{ws} was higher than the specified prior with the posterior median of 0.26. It indicates that the crop started experiencing water stress at about 52% of AWHC. The management allowable depletion of is reported to be 55% for barley and oats which indicates that these species start experiencing water stress when 55% of the AWHC is depleted (Datta et al., 2017).

4.4. Parameters from the hierarchical component of the model

The hierarchical stochastic component allowed us to account for random effects in the system, the plot effect in this case, thus decomposing the residual error. In this case the plot effects for LAI, biomass, and yield were small, but in other cases the model structure outlined here could be readily extended to allow for modeling random block or location effects that are not modeled in a mechanistic manner (e.g. site-specific disease pressure and pest infestation). The negative bias in the plot effect for soil moisture (Table 3) indicates that the model should be improved to better characterize the changes in soil moisture. The low value of NSE for PAW also points towards a potential model bias, whereas the relatively high value of d suggests that the model is accurate in predicting the overall trend (Fig. 2).

5. Model limitations

The model we presented is a simple minimalist crop growth model and cannot mechanistically accommodate the impact of different management practices such as fertilizer application and tillage. However, the hierarchical component of the model would still be able to account for the effects of different treatments. In addition, the model is presented in the form of a mathematical system rather than as software. In its current form, modelers who have experience with R (R Core Team, 2020) and Stan (Stan Development Team, 2020b) will be able to utilize the model and modify it to fit their experimental design.

6. Summary and conclusions

The methodology we presented in this article can be successfully used to estimate crop growth parameters and to understand system dynamics of a winter wheat cropping system. The model predicted LAI and biomass with more success than PAW and yield. Model evaluation statistics indicated overall satisfactory performance of the +WB model, however, potential approaches for further improvements of the model were identified such as accounting for carbohydrate remobilization during grain filling and specifying separate temperature functions for RUE_p and phenology. The biologically meaningful parameters allow crop scientists to make direct physiological interpretations. Estimating parameters along with their uncertainties provides a fuller picture of accuracy and reliability of the estimates. This approach of modeling allows mechanistic representation of physiological processes while also accounting for experimental design factors (such as plot within block effects) that are not typically represented within purely deterministic models. Thus, this modeling approach can be adapted to specific needs and circumstances of a given project.

CRedit authorship contribution statement

Pratishtha Poudel: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Phillip D. Alderman:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – review & editing. **Tyson E. Ochsner:** Methodology, Writing – review & editing. **Romulo P. Lollato:** Data curation, Writing – review & editing.

Declaration of competing interest

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.

Data availability

Data will be made available on request.

Acknowledgments

Special thanks to Donald Riffel for his help uncovering several important typographical errors in the model equations. This research was supported by the National Science Foundation, United States under Grant No. OIA-1826820.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.compag.2024.108618>.

References

- Albrizio, R., Steduto, P., 2005. Resource use efficiency of field-grown sunflower, sorghum, wheat and chickpea: I. Radiation use efficiency. *Agricult. Forest Meteorol.* 130 (3–4), 254–268.
- Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. FAO Irrigation and Drainage Paper No. 56, Vol. 56, No. 97. Food and Agriculture Organization of the United Nations, Rome, e156.
- Amir, J., Sinclair, T.R., 1991. A model of water limitation on spring wheat growth and yield. *Field Crops Res.* 28 (1–2), 59–69.
- Asseng, S., Zhu, Y., Basso, B., Wilson, T., Cammarano, D., 2014. Simulation modeling: Applications in cropping systems. In: Van Alfen, N.K. (Ed.), *Encyclopedia of Agriculture and Food Systems*. Academic Press, Oxford, pp. 102–112. <https://dx.doi.org/10.1016/B978-0-444-52512-3.00233-3>, URL: <https://www.sciencedirect.com/science/article/pii/B9780444525123002333>.
- Attia, A., Rajan, N., Xue, Q., Nair, S., Ibrahim, A., Hays, D., 2016. Application of DSSAT-CERES-Wheat model to simulate winter wheat response to irrigation management in the Texas High Plains. *Agricult. Water Manag.* 165, 50–60.
- Auguie, B., 2017. gridExtra: Miscellaneous functions for "Grid" graphics. URL: <https://CRAN.R-project.org/package=gridExtra>. R package version 2.3.
- Bechini, L., Bocchi, S., Maggiore, T., Confalonieri, R., 2006. Parameterization of a crop growth and development simulation model at sub-model components level. An example for winter wheat (*Triticum aestivum* L.). *Environ. Model. Softw.* 21 (7), 1042–1054.
- Berhe, A.A., Kisekka, I., Prasad, P.V.V., Holman, J., Foster, A.J., Lollato, R., 2017. Assessing wheat yield, biomass, and water productivity responses to growth stage based irrigation water allocation. *Trans. ASABE* 60 (1), 107–121.
- Betancourt, M., 2017. A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.
- Betancourt, M., Girolami, M., 2015. Hamiltonian Monte Carlo for hierarchical models. *Curr. Trends Bayesian Methodol. Appl.* 79 (30), 2–4.
- Caviglia, O.P., Sadras, V.O., Andrade, F.H., 2004. Intensification of agriculture in the south-eastern Pampas: I. Capture and efficiency in the use of water and radiation in double-cropped wheat–soybean. *Field Crops Res.* 87 (2–3), 117–129.
- Crofts, H.J., 1989. On defining a winter wheat. *Euphytica* 44 (3), 225–234.
- Datta, S., Taghvaeian, S., Stivers, J., 2017. Understanding Soil Water Content and Thresholds for Irrigation Management. Technical Report, Oklahoma Cooperative Extension Service.
- Davidson, D.J., Chevalier, P.M., 1992. Storage and remobilization of water-soluble carbohydrates in stems of spring wheat. *Crop Sci.* 32 (1), 186–190.
- Hargreaves, G.H., Samani, Z.A., 1982. Estimating potential evapotranspiration. *J. Irrigation Drainage Division* 108 (3), 225–230.
- Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. *Eur. J. Agron.* 18 (3–4), 235–265.
- Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., L, M.R., M, F.D., J, S.C., 2003. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18 (3–4), 267–288.
- Ko, J., Piccinini, G., Marek, T., Howell, T., 2009. Determination of growth-stage-specific crop coefficients (Kc) of cotton and wheat. *Agricult. Water Manag.* 96 (12), 1691–1697.
- Kukul, M.S., Irmak, S., 2020. Light interactions, use and efficiency in row crop canopies under optimal growth conditions. *Agricult. Forest Meteorol.* 284, 107887.
- Lamsal, A., Welch, S.M., White, J.W., Thorp, K.R., Bello, N.M., 2018. Estimating parametric phenotypes that determine anthesis date in *Zea mays*: Challenges in combining ecophysiological models with genetics. *PLoS One* 13 (4), e0195841.
- Li, G., Yu, M., Fang, T., Cao, S., Carver, B.F., Yan, L., 2013. Vernalization requirement duration in winter wheat is controlled by T a VRN-A 1 at the protein level. *Plant J.* 76 (5), 742–753.
- Lollato, R.P., Bavia, G.P., Perin, V., Knapp, M., Santos, E.A., Patignani, A., DeWolf, E.D., 2020. Climate-risk assessment for winter wheat using long-term weather data. *Agron. J.* 112 (3), 2132–2151.
- Lollato, R.P., Edwards, J.T., 2015. Maximum attainable wheat yield and resource-use efficiency in the southern Great Plains. *Crop Sci.* 55 (6), 2863–2876.
- Lollato, R.P., Edwards, J.T., Ochsner, T.E., 2017. Meteorological limits to winter wheat productivity in the US southern Great Plains. *Field Crops Res.* 203, 212–226.
- Mascheroni, P., Savopoulos, S., Alfonso, J.C.L., Meyer-Hermann, M., Hatzikirou, H., 2021. Improving personalized tumor growth predictions using a Bayesian combination of mechanistic modeling and machine learning. *Commun. Med.* 1, 19.
- Maulana, F., Anderson, J.D., Butler, T.J., Ma, X.-F., 2019. Improving dual-purpose winter wheat in the southern Great Plains of the United States. In: *Recent Advances in Grain Crops Research*. IntechOpen, pp. 1–16.
- McMaster, G.S., Edmunds, D.A., Marquez, R., Haley, S., Buchleiter, G., Byrne, P., Green, T.R., Erskine, R., Lighthart, N., Kipka, H., Fox, F., Wagner, L., Tatarko, M.M., Ascough, II, J., 2019. Winter wheat phenology simulations improve when adding responses to water stress. *Agron. J.* 111 (5), 2350–2360.
- McPherson, R.A., Fiebrich, C.A., Crawford, K.C., Kilby, J.R., Grimsley, D.L., Martinez, J.E., Basara, J.B., Illston, B.G., Morris, D.A., Kloesel, K.A., Melvin, A.D., Shrivastava, H., Wolfenbarger, J.M., Bostic, J.P., Demko, D.B., Elliott, R.L., Stadler, S.J., Carlson, J.D., Sutherland, A.J., 2007. Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *J. Atmos. Ocean. Technol.* 24 (3), 301–321.
- Meredith, M., Kruschke, J., 2018. HDInterval: Highest (Posterior) density intervals. URL: <https://CRAN.R-project.org/package=HDInterval>. R package version 0.2.0.
- Monsi, M., Saeki, T., 2005. On the factor light in plant communities and its importance for matter production. *Ann. Botany* 95 (3), 549.
- Monteith, J.L., 1965. Evaporation and environment. In: *Symposia of the Society for Experimental Biology*, Vol. 19. Cambridge University Press (CUP) Cambridge, pp. 205–234.
- Munaro, L.B., Hefley, T.J., DeWolf, E., Haley, S., Fritz, A.K., Zhang, G., Haag, L.A., Schlegel, A.J., Edwards, J.T., Marburger, D., Alderman, P., Jones-Diamond, S.M., Johnson, J., Lingenfelter, J.E., Uneda-Trevisoli, S.H., Lollato, R.P., 2020. Exploring long-term variety performance trials to improve environment-specific genotype × management recommendations: A case-study for winter wheat. *Field Crops Res.* 255, 107848.
- Muurinen, S., Peltonen-Sainio, P., 2006. Radiation-use efficiency of modern and old spring cereal cultivars and its response to nitrogen in northern growing conditions. *Field Crops Res.* 96 (2–3), 363–373.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* 10 (3), 282–290.
- Nielsen, D.C., Miceli-Garcia, J.J., Lyon, D.J., 2012. Canopy cover and leaf area index relationships for wheat, triticale, and corn. *Agron. J.* 104 (6), 1569–1573.
- Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* 193 (1032), 120–145.
- Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. *Eur. J. Agron.* 10 (1), 23–36.
- Poudel, P., Bello, N.M., Lollato, R.P., Alderman, P.D., 2022. A hierarchical Bayesian approach to dynamic ordinary differential equations modeling for repeated measures data on wheat growth. *Field Crops Res.* 283, 108549. <http://dx.doi.org/10.1016/j.fcr.2022.108549>, URL: <https://www.sciencedirect.com/science/article/pii/S0378429022001204>.
- Pradhan, S., Sehgal, V.K., Bandyopadhyay, K.K., Panigrahi, P., Parihar, C.M., Jat, S.L., 2018. Radiation interception, extinction coefficient and use efficiency of wheat crop at various irrigation and nitrogen levels in a semi-arid location. *Indian J. Plant Physiol.* 23 (3), 416–425.
- Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.* 100 (2), 81–92.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: <https://www.R-project.org/>.
- Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Physics* 1 (5), 318–333.
- Ritchie, J.T., 1991. Wheat phasic development. *Modeling Plant Soil Syst.* 31, 31–54.
- Robertson, M.J., Brooking, I.R., Ritchie, J.T., 1996. Temperature response of vernalization in wheat: modelling the effect on the final number of mainstem leaves. *Ann. Botany* 78 (3), 371–381.
- Rodriguez, D., Keltjens, W.G., Goudriaan, J., 1998. Plant leaf area expansion and assimilate production in wheat (*Triticum aestivum* L.) growing under low phosphorus conditions. *Plant Soil* 200 (2), 227–240.

- Schad, D.J., Betancourt, M., Vasisht, S., 2019. Toward a principled Bayesian workflow in cognitive science. *arXiv preprint arXiv:1904.12765*.
- Sciarresi, C., Patrignani, A., Soltani, A., Sinclair, T., Lollato, R.P., 2019. Plant traits to increase winter wheat yield in semiarid and subhumid environments. *Agron. J.* 111 (4), 1728–1740.
- Slafer, G.A., Rawson, H.M., 1994. Sensitivity of wheat phasic development to major environmental factors: A re-examination of some assumptions made by physiologists and modellers. *Funct. Plant Biol.* 21 (4), 393–426.
- Stan Development Team, 2020a. CmdStan user's guide, version 2.25.0. URL: <http://mc-stan.org>.
- Stan Development Team, 2020b. Stan modeling language users guide and reference manual, version 2.25.0. URL: <http://mc-stan.org>.
- Steduto, P., Hsiao, T.C., Raes, D., Fereres, E., 2009. AquaCrop—The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. *Agron. J.* 101 (3), 426–437.
- Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. *Eur. J. Agron.* 18 (3–4), 289–307.
- Tenreiro, T.R., García-Vila, M., Gómez, J.A., Jimenez-Berni, J.A., Fereres, E., 2020. Water modelling approaches and opportunities to simulate spatial water variations at crop field level. *Agricult. Water Manag.* 240, 106254.
- United States Department of Agriculture, Soil Conservation Service, Engineering Division, 1986. Urban Hydrology for Small Watersheds, No. 55. Engineering Division, Soil Conservation Service, US Department of Agriculture.
- Van Diepen, C.A.v., Wolf, J.v., Van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. *Soil Use Manag.* 5 (1), 16–24.
- Wagle, P., Gowda, P.H., Northup, B.K., Neel, J.P.S., Starks, P.J., Turner, K.E., Moriasi, D.N., Xiao, X., Steiner, J.L., 2021. Carbon dioxide and water vapor fluxes of multi-purpose winter wheat production systems in the US southern Great Plains. *Agricult. Forest Meteorol.* 310, 108631.
- Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R.P., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., et al., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nature Plants* 3 (8), 1–13.
- Wang, E., Zhao, Z., 2013. Improving APSIM for simulation of temperature response of wheat (APSIM-WheatE). *Model. Wheat Response High Temperature* 29.
- Wickham, H., 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, URL: <https://ggplot2.tidyverse.org>.
- Wickham, H., 2017. tidyverse: Easily install and load the 'Tidyverse'. URL: <https://CRAN.R-project.org/package=tidyverse>. R package version 1.2.1.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. *J. Open Source Softw.* 4 (43), 1686. <http://dx.doi.org/10.21105/joss.01686>.
- Williams, J.R., 1991. Runoff and water erosion. In: *Modeling Plant and Soil Systems*. John Wiley & Sons, Ltd, pp. 439–455. <http://dx.doi.org/10.2134/agronmonogr31.c18>, URL: <https://acsess.onlinelibrary.wiley.com/doi/abs/10.2134/agronmonogr31.c18>.
- Willmott, C.J., 1981. On the validation of models. *Phys. Geogr.* 2 (2), 184–194.
- Xie, Y., 2020. knitr: A general-purpose package for dynamic report generation in R. URL: <https://yihui.org/knitr/>. R package version 1.29.
- Yang, R., Wang, Z., Chen, J., 2021. An integrated approach of mechanistic-modeling and machine-learning for thickness optimization of frozen microwaveable foods. *Foods* 10 (4), 763.
- Yang, J., Zhang, J., Huang, Z., Zhu, Q., Wang, L., 2000. Remobilization of carbon reserves is improved by controlled soil-drying during grain filling of wheat. *Crop Sci.* 40 (6), 1645–1655.
- Zhu, H., 2019. kableExtra: Construct complex table with 'kable' and pipe syntax. URL: <https://CRAN.R-project.org/package=kableExtra>. R package version 1.1.0.