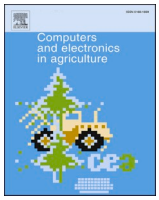
Computers and Electronics in Agriculture 227 (2024) 109513

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



Enhancing crop model parameter estimation across computing environments: Utilizing the GLUE method and parallel computing for determining genetic coefficients

Thiago Berton Ferreira [a,*], Vakhtang Shelia [a,b], Cheryl Porter [a], Patricia Moreno Cadena [c], Montserrat Salmeron Cortasa [d], Muhammad Sohail Khan [a,e], Willingthon Pavan [a], Gerrit Hoogenboom [a,b]

[a] Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, USA
[b] Global Food Systems Institute, University of Florida, Gainesville, FL, USA
[c] The Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), Nairobi, Kenya
[d] Department of Plant and Soil Sciences, College of Agriculture Food and Environment, University of Kentucky, Lexington, KY, USA
[e] Department of Horticulture, Faculty of Agriculture, Gomal University, Dera Ismail Khan, Khyber Pakhtunkhwa, Pakistan

ARTICLE INFO

Keywords:
Genetic-specific parameters
DSSAT
Cropping System Model
Ecotype parameters
High-Performance Computing environments

ABSTRACT

Estimating genetic coefficients is essential to accurately simulate crop development and growth for modeling studies but has been challenging due to lack of robust and fast procedures. While there are several optimization techniques, the Generalized Likelihood Uncertainty Estimation (GLUE) is a Bayesian method that is popular among the modeling community due to its application for sensitivity and uncertainty analysis and capability to explore the global parameter space. However, the time required for its search method to estimate the optimal parameter set is a significant constraint and limitation. Parallel computing has emerged as a solution to boost the efficiency of genetic coefficient calibration using GLUE. In this study, we introduce a new system that leverages parallel computing for calibrating genetic inputs for crop growth models within the Cropping System Model (CSM) of the Decision Support System for Agrotechnology Transfer (DSSAT). Designed and tested for both conventional and High-Performance Computing (HPC) environments, the Generalized Likelihood Uncertainty Estimation Parallelized (GLUEP) is available for most crops that are simulated with DSSAT-CSM and provides a user-friendly graphical interface within the DSSAT software. It accelerates the genetic-specific parameter calibration process and adds new functionality that enables users to optimize intrinsic model parameters, which were previously unavailable for calibration purposes. Four case studies using cultivars for wheat, maize, soybean, and potato showcase the application of GLUEP. We also conducted a comparison with DSSAT-GLUE and evaluated the performance gains of GLUEP for multiple operational systems, including Windows, MacOS, and Linux, as well as under conventional and HPC environments. The multi-core processing results indicate performance improvements across all computer systems that were analyzed. The comparison between the sequential processing of DSSAT-GLUE and the parallel processing of GLUEP indicates a reduction in execution time ranging from 87.4% to 95.4%. These results highlight the GLUEP capabilities in streamlining the calibration process, enabling more efficient and accurate predictions for crop growth modeling studies.

1. Introduction

Accurate representation of real-world processes through computer models fundamentally requires model parameterization and calibration (Martre et al., 2015; Alderman and Stanfill, 2017). The simulation of biological and environmental interactions using process-based crop growth models involves estimating model parameters related to crop phenological and biophysical response to the agricultural ecosystem (Ahmed et al., 2020; Wallach et al., 2021a). Therefore, simulating complex interactions associated with plant growth and development require a reliable method to ensure the accuracy of model predictions as different genotypes within the same species require specific calibrations

* Corresponding author at: Frazier Rogers Hall, PO Box 110570, Gainesville, FL 32611-0570, USA.
E-mail address: t.berton@ufl.edu (T. Berton Ferreira).

https://doi.org/10.1016/j.compag.2024.109513
Received 22 May 2024; Received in revised form 6 September 2024; Accepted 27 September 2024

(Jha et al., 2022). Although crop growth models vary according to the set of equations and general structure, the calibration of parameters is essentially dependent on the field-collected data related to the target population (White and Hoogenboom, 1996; Hunt and Boote, 1998; Wallach et al., 2023). Field observations detailing information on crop phenology, such as anthesis and maturity dates, and dry canopy growth, including aboveground biomass, leaf area index, and final yield, are essential for estimating the genetic coefficients of a given crop cultivar (Boote et al., 2001). Model calibration is a required step for every crop growth model to evaluate the model's capability to reproduce real-world observations (Boote et al., 2021). The Cropping System Model (CSM) of the Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003; Hoogenboom et al., 2019; Hoogenboom et al., 2023) uses genetic-specific parameters (GSPs) to simulate the crop responses to environmental conditions such as temperature, solar radiation, and rainfall. More specifically, the ecotype file (FILEE), contains crop genetic coefficients that define characteristics often similar across multiple cultivars, while the cultivar file (FILEC) defines coefficients unique to a particular crop cultivar, for example, photothermal days to flowering and sensitivity to photoperiod. Hence, GSPs in the DSSAT-CSM refer to genetic coefficients both in the cultivar and ecotype files.

Several tools have been developed to calculate and estimate the optimal set of DSSAT GSPs for representing crop genotypes according to data observations (Hunt et al., 1993; He et al., 2010; Gao et al., 2020; Hyun et al., 2022; Memic et al., 2021). These cultivar calibration tools conduct parameter estimation through statistical analysis and inference, comparing with field measurements to evaluate the results. The Generalized Likelihood Uncertainty Estimation (GLUE) method is a common approach used for calibrating parameters (Beven and Binley, 2014; Mirzaei et al., 2015; Li et al., 2020; Ragab et al., 2020) that has been implemented for a few crop growth models (Buddhaboon et al., 2018; Li et al., 2018; Sheng et al., 2019). Originally developed to optimize the parameter set of hydrological models (Beven and Binley, 1992), the GLUE method was adapted to estimate crop model parameters within the DSSAT-CSM (He et al., 2010).

In addition to GLUE, there are other methods for calibrating cultivar-specific parameters. Hunt et al. (1993) developed one of the first tools for estimating cultivar coefficients for crop models within the DSSAT-CSM called the Genotype Coefficient Calculator (GENCALC). The software initializes the calibration with approximate coefficients. GENCALC fine-tunes the model parameters through iterative simulations, comparing them with the observed crop measurements and then altering each parameter individually to better fit the observations (Anothai et al., 2008). This stepwise method can adjust the genetic coefficients but requires more manual operations than GLUE (Buddhaboon et al., 2018). Gao et al. (2020) conducted a comparative study on the performance of GLUE, Markov Chain Monte Carlo (MCMC), and Ordinary Least Square (OLS) methods to calibrate rice phenology using the DSSAT-CSM-CERES-Rice model. For calibrating phenology, their findings indicate that OLS is the fastest calibration approach and MCMC is a promising method to quantify prediction uncertainty. GLUE had a slower computational performance but reduced variance in model error variance and less parameter uncertainty compared to the other methods evaluated by Gao et al. (2020).

Gradient-based algorithms can also be applied for parameter estimation, using prior information from model calibration to identify the most likely search directions (Akhavizadegan et al., 2021; Shahhosseini et al., 2021). An example of that is the Parameter ESTimation model (PEST), which can be used to calibrate parameters of computer models in general, including crop growth models (Doherty et al., 2010). Although calibrating genetic coefficients through gradient-based algorithms can reduce execution time, this method is more likely to reach a local minimum (Wallach et al., 2021a). The convergence to the local minimum has been observed in different crop growth models within APSIM and DSSAT when using PEST (Harrison et al., 2019; Ma et al., 2020). Meanwhile, global search algorithms, such as GLUE, can avoid convergence to a local optimum and typically demand more iterations. However, this Bayesian approach has previously shown effectiveness in calibrating parameters for different crops and cropping system models such as DSSAT, APSIM, and a simplified version of STICS (Tremblay and Wallach, 2004; Li et al., 2018; Sheng et al., 2019; Afzal et al., 2024; Dahri et al., 2024).

In recent efforts, Memic et al. (2021) introduced a system for calibrating cultivar genetic coefficients of DSSAT crop models using time series data, measuring Goodness-of-Fit based on the root mean square error variances. This method can potentially provide an advantage compared to DSSAT-GLUE, which considers only the end-of-season data for model calibration. However, as more data observations are considered, the time required for conducting such calibration can increase exponentially compared to the GLUE method.

To enhance the execution efficiency of GLUE in DSSAT, Hyun et al. (2022) developed a system named Generalized Likelihood Uncertainty Estimation Orchestration System (GLUEOS), using Kubernetes. The system orchestrates multiple instances of GLUE containers and parallelizes the cultivar calibration process. Software layers written in C++ for handling the communication between containers and a Python-based web interface were developed on top of the original R code of DSSAT GLUE. The introduction of additional layers to the application increases the integration complexity and code maintenance. While GLUEOS demonstrated efficiency on single board and high-end desktop computers, the authors did not aim for compatibility with High-Performance Computing (HPC) clusters, where parallelized systems could substantially enhance computational efficiency and reduce computational time. Furthermore, the system was developed for a single crop growth model in the DSSAT-CSM and was not generally designed for expanding the cropping system environment of DSSAT for other crops.

This paper aims to introduce a novel software for genetic coefficient calibration called the Generalized Likelihood Uncertainty Estimation Parallelized (GLUEP). By harnessing computer parallelism, GLUEP offers a faster and more efficient method to expand and explore the crop model hyperparameter space, allowing a more robust and reliable GSP calibration. The study objectives are to (i) provide the basic concepts of the algorithm and parallelism in GLUEP, (ii) describe the software architecture and components, (iii) demonstrate through case studies the functionalities and integration with DSSAT-CSM, (iv) demonstrate system versatility and performance across different computer platforms and computing environments.

2. Materials and methods

2.1. Generalized Likelihood Uncertainty Estimation

The Generalized Likelihood Uncertainty Estimation is initialized by defining a prior distribution based on each parameter's defined lower and upper limits (He et al., 2010). These limits delineate a range of possible values that a genetic parameter can assume. By default, the GLUE for genetic coefficient calibration of crop model parameters in DSSAT assumes a uniform probability distribution for all parameter prior distribution. Values are randomly selected from the prior distribution and inserted into a vector of size z corresponding to the total number of genetic coefficients that are being calibrated, generating a combination of one potential value for each individual genetic parameter. This process repeats and generates a matrix of z columns and size N corresponding to the number of samples, also known as parameter sets $\{\theta_1, \theta_2, \cdots, \theta_i, \cdots \theta_N\}$. Columns of the resulting matrix correspond to a genetic coefficient and rows to the parameter set.

Once the matrix of random parameter sets is created, model runs are conducted for every generated parameter set (He et al., 2010). The simulation outputs for each of the N parameter vectors θ_i are tabulated alongside the corresponding observations that are defined in the end-of-season experimental data file (fileA). Based on the simulation results and observed data, the likelihood value for each parameter set is calculated

using Gaussian probability density function (Equation (1).

$$L(\theta_i|O) = \prod_{j=1}^{M} \frac{1}{\sqrt{2\pi\sigma_o^2}} \exp\left(-\frac{\left(O_j - P_j(\theta_i)\right)^2}{2\sigma_o^2} \right) \quad (1)$$

where $L(\theta_i|O)$ is the likelihood value of parameter set i, given the set of observations,

M is the number of observations,

σ_o^2 is the variance of model error,

O_j is the jth observation of O,

$P_j(\theta_i)$ is the jth simulated model output under parameter set θ_i.

The probability of each parameter set is then computed following the equation:

$$p(\theta_i) = \frac{L(\theta_i|O)}{\sum_{i=1}^{N} L(\theta_i|O)} \quad (2)$$

where $p(\theta_i)$ is the probability of the ith parameter set θ_i.

A posterior distribution is then generated for each parameter set. Based on that, GLUE searches for the parameter set with the highest probability value within the N model runs conducted for the genetic coefficient calibration. Under this context, a practical problem using the GLUE method is that the simulation quality in reproducing the observed values is likely to be poor for a small number of N. For instance, the DSSAT-GLUE guidelines recommend using at least 6000 for N (Gao et al., 2020). Under the Monte Carlo sampling perspective, increasing the magnitude of N expands the parameter sets and enhances the likeliness of finding the optimal parameter. However, this leads to a substantial increase in the execution time for parameter calibration. Hence, there is a pressing need to develop a new tool to accelerate the GSPs calibration, enabling users to attain the optimal parameter set in less time.

2.2. Software architecture

The GLUEP was developed to calibrate genetic-specific parameters using the GLUE method and computer parallelism to streamline execution time and improve system performance. Developed in R, GLUEP allows a seamless deployment aligned to the R-based architecture of DSSAT-GLUE. The built-in R package "parallel" was used to develop the parallel processing system. As each individual model execution uses different parameter sets, they are independent and unrelated to each other. Therefore, GLUEP employs the same mathematical approach defined in the GLUE method, but with the parameter sets partitioned into chunks, allowing multiple CPUs to process the simulations concurrently. As a result, the time required for calibrating genetic coefficients can be significantly reduced.

The GLUEP source code is publicly available on the DSSAT Foundation's GitHub at https://github.com/DSSAT/GLUEP (GLUEP Source code, 2024) and integrated into the DSSAT platform version 4.8.5 (Hoogenboom et al., 2024). The system features a graphical user interface in Delphi similar to the DSSAT-GLUE. The GLUEP source code optimizes the calibration process by reducing the volume of the crop model's output files and eliminating unnecessary outputs in the system's terminal, an aspect overlooked in DSSAT-GLUE. The interface generates a CSV file as input and executes the GLUEP R program code. The system automatically fragments the workload across the available computer cores and orchestrates the execution across the CPUs. The system uses communication via network sockets between a master process, which orchestrates the calibration, and worker processes (R Core Team, 2024) that perform the calibration. This communication method makes GLUEP a cross-platform software, compatible with any operational system (e.g., Windows, MacOS, Unix-like systems), as the software relies on the computer network for data transmission across the nodes. The master node automatically fragments the workload (e.g., the total parameter set) into roughly equal-sized chunks based on the number of runs and allocated computer nodes for the calibration (R Core Team, 2024). Multithreading is often associated with parallel processing, where a single CPU can execute two or more instructions simultaneously, also known as threads. However, by default, R internal structure is single-threaded, meaning each CPU functions as a single core and workers cannot be divided to the thread-level (Ramalakshmi and Kompala, 2017). When a worker completes a task, a response is sent back to the master process, which continuously monitors and waits for the cluster of workers to finish. Once all parameter sets are tested, the simulation outputs are unified into a file for generating the posterior distribution, and GLUEP identifies the optimal parameter set for the given genotype as calibration process output (Fig. 1). For HPC environments, the

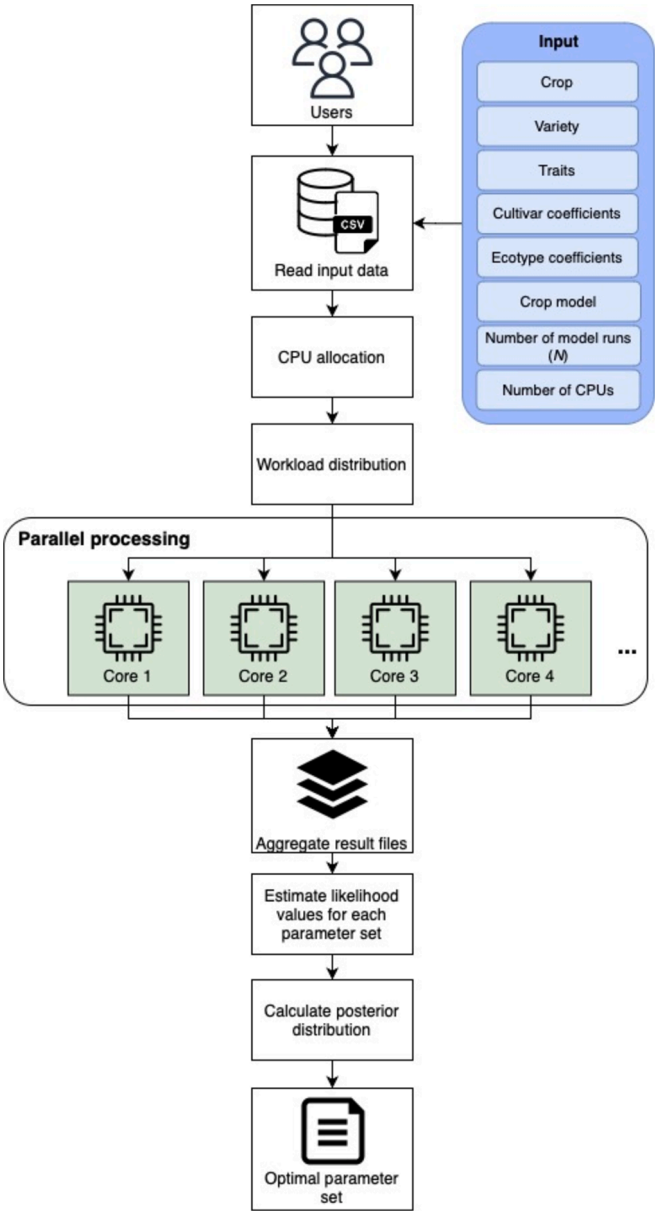


Fig. 1. Conceptual diagram and process flow of the Generalized Likelihood Uncertainty Estimation Parallelized (GLUEP) for crop genetic coefficient estimation through parallel computing.

workload is handled through job schedulers, such as the Simple Linux Utility for Resource Management (SLURM), which manages and allocates resources for executing computational tasks, also known as jobs, and handles the job queue. SLURM is one of the most utilized open-source job scheduling systems, and GLUEP autonomously extracts the number of processors requested within SLURM job scripts for calibration of the GSPs.

GLUEP also expands the parameter calibration capabilities of DSSAT models by including an additional feature for calibrating ecotype-related genetic coefficients, a function not available in DSSAT-GLUE or similar tools that have previously been developed. These crop model parameters are essential for describing genetic characteristics in DSSAT-CSM models. Unlike the cultivar-specific coefficients, these generic parameters often represent multiple cultivars, such as genetically related varieties. This newly introduced feature allows users to estimate coefficients associated with common genetic traits across a group of cultivars. Incorporating and calibrating more genetic coefficients, essential for driving crop model simulations, expands the capabilities of exploring the parameter space. Consequently, a more thorough search for the optimal parameter set will likely result in a model capable of reproducing data observations with increased precision and reliability.

2.3. Case studies

The application and functionality of this parallelized computing approach for estimating crop genetic coefficients are illustrated using four case studies. Each case exemplifies the calibration of cultivars for four different crops and four different crop growth models within the DSSAT CSM based on data from field experiments. The first case study showcases the calibration of the wheat (*Triticum aestivum L.*) cultivar MGS-Brilhante using the CSM-CERES-Wheat model. The second case study highlights the calibration of the potato (*Solanum tuberosum L.*) cultivar Arizona using the CSM-SUBSTOR-Potato model. Both wheat and potato cultivars were part of newly introduced experiments in DSSAT and have not been previously calibrated with these models.

The third case study comprises the calibration of the soybean (*Glycine* max *L.*) cultivar S80-P2 using the CSM-CROPGRO-Soybean model. Bao et al. (2015) previously calibrated this cultivar for DSSAT Version 4.5, but the complete parameter set was never released.

The fourth case study covers the calibration of the maize (*Zea mays L.*) hybrid DeKalb 52–59 AF2 (VT3) for genetic coefficients within the cultivar and ecotype files using the CSM-CERES-Maize model. Anothai et al. (2013) had previously calibrated the genetic coefficients within the cultivar file for this maize hybrid and reported the post-calibration parameter set. The results from this earlier calibration were included for comparison to assess the effectiveness of the new feature for calibrating both cultivar and ecotype GSPs.

2.3.1. Case study 1: GSP calibration of a spring wheat cultivar

Data for a spring wheat cultivar MGS-BRILHANTE were obtained from a series of field experiments conducted for seven years (2013 to 2019) at the experimental station of the Agricultural Research Company of Minas Gerais (EPAMIG) in Patos de Minas, Minas Gerais, Brazil (18°31′04.0″S, 46°26′25.0″W). The field trials were originally designed to evaluate the efficiency of fungicides against wheat blast (*Magnaporthe oryzae* pathotype *Triticum*). The experimental plots were established uniformly, each with an area of 5 m^2 and organized in five rows with 0.2 m of separation between them. For every year and each experiment, a set of 12 sowing dates were selected with intervals of 7 to 10 days between each sowing. During the trials, it was observed that some experimental plots from 2013 to 2018 were not affected by the fungal disease (Ascari et al., 2021).

Soil profile data were extracted from a nearby location in Presidente Olegário, Minas Gerais, Brazil (18°07′16.5″S 46°29′20.5″W) and weather data were obtained from the Brazilian daily weather gridded data for 2013 to 2019 (Xavier et al., 2022). Crop management was defined according to recommendations from the Brazilian Agricultural Research Corporation (EMBPAPA) for the region (Chargas et al., 2020). The observed data includes maturity date for calibrating genetic coefficients related to phenology and yield for calibrating crop growth genetic coefficients.

Based on the disease-free data of this six-year experiment, the genetic coefficients only from the cultivar file of the MGS-BRILHANTE cultivar were used for calibration (Table 1). The cultivar calibration was conducted based on four years of data (2013, 2014, 2015, and 2017) with $N = 100{,}000$. The remaining two years of field observed data (2016 and 2018) were used for calibrated model evaluation.

2.3.2. Case study 2: GSP calibration of a potato cultivar

A series of field experiments spanning two years were conducted under sub-tropical growing conditions of Dera Ismail Khan, Khyber Pakhtunkhwa, Pakistan (between 31°43′ N latitude, 70°49′ E longitude, and 177 m altitude) during the winter seasons (October-March) of 2017–18 and 2018–19. The experiments were set up in a randomized complete-block design (RCBD) with two factors (planting date and genotype) and three replications. The potato cultivar Arizona was planted on four planting dates, i.e., very-early: 02 Oct (*T1*), early: 14 Oct (*T2*), late: 26 Oct (*T3*), and very-late: 07 Nov (*T4*). Thus, plants were deliberately exposed to seasonal weather fluctuations caused by the different

Table 1
Description of the wheat genetic coefficients of the CSM-CERES-Wheat crop model.

Parameter	Description	Trait	Range	Unit
P1V	Optimum vernalizing temperature	Phenology	0–60	°C day^{-1}
P1D	Photoperiod response	Phenology	0–200	% reduction in rate/10 h^{-1} drop in photoperiod
P5	Grain filling phase duration	Phenology	100–999	°C day^{-1}
G1	Kernel number per unit canopy weight at anthesis	Dry matter growth	10–50	# g
G2	Kernel size under optimum conditions	Dry matter growth	10–80	mg
G3	Non-stressed mature tiller weight	Dry matter growth	0.5–8	g dry weight
PHINT	Interval between successive leaf tip appearances	–	30–150	°C day^{-1}

Table 2
Description of the potato genetic coefficients in the CSM-SUBSTOR-Potato crop model.

Parameter	Description	Trait	Range	Unit
P2	Tuber initiation sensitivity to photoperiod	Phenology	0.2–0.9	Dimensionless
TC	Upper critical temperature for tuber initiation	Phenology	15–22	°C
G2	Leaf area expansion rate after tuber initiation	Dry matter growth	900–2100	cm^2 m^{-2} day^{-1}
G3	Potential tuber growth rate	Dry matter growth	21–26	g m^{-2} day^{-1}
PD	Index that suppresses tuber growth	Dry matter growth	0.5–1	Dimensionless

planting dates. Each plot (16.2 m^2) had six ridges with 12 plants each. Well-sprouted, healthy, and uniform-sized seed tubers of the potato genotypes were planted at 10 cm depth on the ridge, one seed per hole and 0.3 m apart. A gap of 0.75 m was maintained between any two ridges. Standard practices were followed during the planting of the seed tubers. Local recommended cultural practices were followed under irrigated conditions to sustain normal crop growth without biotic and abiotic stress.

Detailed observations on crop phenology, growth, and yield-attributing parameters were collected weekly throughout the crop cycle. Also, daily weather inputs (solar radiation (MJ m^{-2}), maximum and minimum temperature (°C), and precipitation (mm), soil profile characteristics, initial soil condition, and field management practices were elaborately collected. Crop measurements made during the first growing season (2017–18) for all the aforementioned four planting dates were used for SUBSTOR-Potato model calibration. Based on these observations, genotype-specific model-input parameters (Table 2) were estimated for the potato genotype through the GLUEP with a total of 300,000 model runs. After the completion of model calibration, the data sets from the second growing season (2018–19) were used to perform model evaluation and to assess the implications of seasonal variation and genotypes on the tuber yield under the sub-tropical growing conditions. For details on the methodology and results, see Khan et al. (2024).

2.3.3. Case study 3: GSP calibration of a soybean cultivar

A series of soybean experiments with six cultivars were conducted in Watkinsville (33°52′N, 83°32′W), Calhoun (34°29′N, 84°58′W), Williamson (33°10′N, 84°24′W), Midville (32°52′N, 82°13′W), Plains (32°2′N, 84°22′W) and Tifton (31° 28′N, 83°31′W), Georgia, United States to quantify the impact of irrigation on soybean growth and development (Bao et al., 2015). Cultivar S80-P2 was selected for this case study since it included a more robust field observation dataset and crop management data for three years (2005 to 2007) with beginning of flowering (R1), maturity date (R8), and seed yield. The soil and weather data are already part of the DSSAT collection of experiments and were extracted from Perkins et al. (1986) and the Georgia Automated Environmental Monitoring Network (AEMN – http://www.georgiaweather.net), respectively. Although the genetic coefficients for the soybean cultivars had been previously calibrated, this data has never been included in the DSSAT repository. Genetic coefficients (Table 3) were calibrated based on the multi-location records for the odd years (2005 and 2007) with $N = 40{,}000$ runs and evaluated with data from 2006.

2.3.4. Case study 4: Calibration of the cultivar and ecotype coefficients for a maize hybrid

Field experiment data from the USDA-ARS Water Management Research Unit experimental site in Greeley, Colorado, United States (40°26′ N, 104°38′ W) were used to calibrate the maize hybrid DeKalb 52–59 AF2 (VT3) (Trout and Bausch, 2012). Anothai et al. (2013) detailed the two-year experimental data (2008 and 2010) as the cultivar coefficients of the maize hybrid (Table 4) was previously calibrated using an earlier version of the CERES-Maize model within DSSAT-CSM Version 4.5. For this case study, the set of genetic coefficients previously obtained was compared with the outcomes of the present study as both cultivar and ecotype coefficients for the DeKalb 52–59 AF2 (VT3) hybrid were recalibrated using the novel functionality of GLUEP. The cultivar calibration process involved the treatments with the highest yield with optimal irrigation conditions of 2008 (100 % of full crop water requirements) and $N = 200{,}000$ runs, while the treatments with optimal irrigation conditions of 2010 and suboptimal irrigation conditions (40 % of full crop water requirements) were used to evaluate the calibrated genetic coefficients.

2.4. Computational performance

A thorough exploration of the parameter space is fundamental for identifying the optimal combination of genetic coefficients that accurately represent a given crop cultivar. A series of analyses was conducted to determine the performance of GLUEP during the GSP calibration process. A comparison of wall clock time between GLUEP and DSSAT-GLUE was included to examine the impact of parallel processing on execution time. The evaluation was conducted for four computing environments with different operational systems (OS) and CPUs (Table 5). The default file paths and system commands in the DSSAT-GLUE source code, originally designed for Windows system, were adapted to support MacOS and Linux environments.

A computational experiment was designed to capture the overall improvement in computational performance with GLUEP. The maize hybrid McCurdy 84aa was previously calibrated for the CSM-CERES-Maize model (Bennett et al., 1989; Hoogenboom et al., 2019) and used for conducting the performance analysis. Six treatments with different combinations of nitrogen and irrigation management were considered. For all computing environments, calibrations were conducted to evaluate performance using $N = 50{,}000$ runs, denoted as a regular volume GSP calibration. Additionally, a calibration employing $N = 200{,}000$ runs was conducted on the HPC to assess the performance of GLUEP with an increased number of runs, referred to as high volume

Table 3
Description of the soybean genetic coefficients of the CSM-CROPGRO-Soybean crop model.

Parameter	Description	Trait	Range	Unit
CSDL	Critical Short Day Length below which reproductive development progresses with no daylength effect	Phenology	11.78–14.6	Hour
PPSEN	Slope of the relative response of development to photoperiod with time	Phenology	0.129–0.385	1 h^{-1}
EM-FL	Time between plant emergence and flower appearance	Phenology	9–28.9	Photothermal Day
FL-SH	Time between first flower and first pod	–	5–10	Photothermal Day
FL-SD	Time between first flower and first seed	Phenology	11–22	Photothermal Day
SD-PM	Time between first seed and physiological maturity	Phenology	22–37.7	Photothermal Day
FL-LF	Time between first flower and end of leaf expansion	–	18–26	Photothermal Day
LFMAX	Maximum leaf photosynthesis rate at 30C, 350 vpm CO_2, and high light	Dry matter growth	1–1.4	mg (CO_2) m^{-2}
SLAVR	Specific leaf area of cultivar under standard growth conditions	Dry matter growth	300–400	cm^2 g^{-1}
SIZLF	Maximum size of full leaf	Dry matter growth	137–230	cm^2
XFRT	Maximum fraction of daily growth that is partitioned to seed and shell	–	1–1	Dimensionless
WTPSD	Maximum weight per seed	Dry matter growth	0.15–0.19	g
SFDUR	Seed filling duration for pod cohort at standard growth conditions	Dry matter growth	17–25.5	Photothermal Day
SDPDV	Average seed per pod under standard growing conditions	Dry matter growth	1.7–2.44	# (seed) pod^{-1}
PODUR	Time required for cultivar to reach final pod load under optimal conditions	–	10–10	Photothermal Day
THRSH	Threshing percentage	–	77–78	%
SDPRO	Fraction protein in seeds	–	0.4–0.405	g (protein) g^{-1} (seed)
SDLIP	Fraction oil in seeds	–	0.2–0.205	g (oil) g^{-1} (seed)

Table 4
Description of the maize genetic coefficients of the CSM-CERES-Maize crop model.

Parameter	Description	Trait	Range	Unit
P1[c]	Thermal time from seedling emergence to the end of the juvenile phase	Phenology	5–450	°C day^{-1}
P2[c]	Extent to which development is delayed for each hour increase in photoperiod above the longest photoperiod at which development proceeds at a maximum rate	Phenology	0–2	% reduction in rate/10 h^{-1} drop in photoperiod
P5[c]	Thermal time from silking to physiological maturity	Phenology	580–999	°C day^{-1}
G2[c]	Maximum possible number of kernels per plant	Dry matter growth	248–990	# g
G3[c]	Kernel filling rate during the linear grain filling stage and under optimum conditions (mg/day)	Dry matter growth	5–16.5	mg
PHINT[c]	Interval between successive leaf tip appearances	–	38–75	°C day^{-1}
TBASE[e]	Base temperature below which no development occurs	Phenology	7–8	°C
TOPT[e]	Temperature at which maximum development rate occurs during vegetative stages	Phenology	33–34	°C
ROPT[e]	Temperature at which maximum development rate occurs for reproductive stages	Phenology	33–34	°C
P20[c]	Day length below which day length does not affect development rate	Phenology	12–12.5	Hour
DJTI[e]	Minimum days from end of juvenile stage to tassel initiation if the cultivar is not photoperiod sensitive	Phenology	3–4	Day
GDDE[c]	Growing degree days per cm seed depth required for emergence	Dry matter growth	5–6	GDD cm^{-1}
DSGFT[e]	GDD from silking to effective grain filling period	Dry matter growth	169–170	GDD
RUE[e]	Radiation use efficiency	Dry matter growth	3.2–4.2	g MJ^{-1}
KCAN[e]	Canopy light extinction coefficient for daily PAR	–	0.75–0.85	Dimensionless

[c] Genetic coefficient defined in the cultivar file (CUL).
[e] Genetic coefficient defined in the ecotype file (ECO).

Table 5
Description of computer resources for GLUEP performance assessment.

Operational System	Computing environment	CPU model	Number of cores per node available
Windows 10	Conventional (Desktop)	Intel Core i5-6500 CPU 3.2 GHz	4
Windows 10	Conventional (Desktop)	Intel Core i7-7700 CPU 3.6 GHz	8
MacOS	Conventional (Laptop)	Apple M1 Pro 8-core	10
Linux Red Hat 8.8 (Ootpa)	High Performance Computing cluster	AMD EPYC 7702 64-Core Processor	128

GSP calibration.

Different numbers of CPUs were used across multiple computing environments, ranging from a single CPU to a maximum of 120 CPUs in the HPC setting. For DSSAT-GLUE, only one CPU core was utilized because the system is a serial program that processes tasks sequentially. Hence, the increase of CPU cores will not improve the DSSAT-GLUE performance. Additionally, the calibration using GLUEP with a single CPU essentially operates as a serial program as there is no parallel execution involved, given the R single-threaded internal structure. Consequently, a comparison between DSSAT-GLUE and GLUEP was conducted to analyze the effectiveness of the optimized code.

Calibrations in conventional/personal computing environments (desktops) were replicated five times. On the other hand, calibrations conducted on the HPC cluster, i.e., the University of Florida HiPerGator, were replicated ten times due to a higher level of uncertainty associated with a shared environment. In total, 300 calibration sets were conducted across all replicates and platforms.

2.5. Evaluation metrics

Statistical analysis was conducted using the R package "ie2misc" and four goodness-of-fit metrics were calculated to evaluate model performance after calibration for each case study. The Root Mean Square Error (*RMSE*) was used to measure the differences between observed and simulated values (Equation (3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T}(M_i - S_i)^2}{T}} \tag{3}$$

where T is the total number of observations,

M_i is the measured value at index i,

S_i is the simulated value at index i.

Willmott's index of agreement (*d-stat*) was used to measure the model prediction error (Willmott et al., 1985). The index values can range from 0 to 1, where a score of 0 indicates no agreement and 1 indicates a perfect match (Equation (4).

$$\text{d-stat} = 1 - \frac{\sum_{i=1}^{T}(M_i - S_i)}{\sum_{i=1}^{T}\left(|M_i - \overline{M}|\right) + \left(|S_i - \overline{M}|\right)}) \tag{4}$$

where $\overline{M}$ is the mean of the measured values.

The r-squared (R^2) measurement was used to determine the proportion of variance that the model can explain (Equation (5).

$$R^2 = 1 - \frac{\sum_{i=1}^{T}(M_i - \widehat{S}_i)^2}{\sum_{i=1}^{T}(M_i - \overline{M})^2} \tag{5}$$

where $\widehat{S}_i$ is the simulated value at index i as given by a linear regression model.

The Nash-Sutcliffe model efficiency coefficient (*NSE*) assessed the model predictions (Nash and Sutcliffe, 1970). *NSE* represents the ratio between RMSE and the variance within the observed values, where a score of 0 indicates a poor predictive skill of the model as it is worse than using the mean of the observed data, while a score of 1 indicates a perfect model response (Equation (6).

$$NSE = 1 - \frac{\sum_{i=1}^{T}(M_i - S_i)^2}{\sum_{i=1}^{T}(M_i - \overline{M})^2} \tag{6}$$

Additionally, metrics to evaluate the performance of parallel algorithms and architectures were also considered. The overall reduction in computing time was measured through speedup, a parallel performance measurement prevenient from Amdahl's law (Grama et al., 1993). Speedup indicates the ratio of time required by a sequential algorithm

(single-threaded) to solve a computational task, known as serial time, to the time a parallel algorithm (multi-threaded) takes to solve the same task, known as parallel time (Equation (7).

$$S(p) = \frac{t_s}{t_p(p)} \tag{7}$$

where $S(p)$ is the speed ratio using p processors,

t_s is the serial execution time required to execute a computational task,

$t_p(p)$ is the parallel execution time required to execute a computational task using p processor(s).

Furthermore, in parallel architectures, each individual CPU tends to experience losses due to the communication and synchronization between cores, adding procedures to execute parallel computing tasks. These additional procedures, regarded as overhead time, ultimately increase the computation time. The parallel efficiency equation, also derived from Amdahl's law, is used to measure this, expressing the speedup ratio to the total number of processors (Equation (8).

$$E(p) = \frac{S(p)}{p} \tag{8}$$

where $E(p)$ is the efficiency of using p processors,

p is the total number of processors.

3. Results

3.1. Case studies

3.1.1. Case study 1: GSP calibration of a spring wheat cultivar

The CSM-CERES-Wheat model with the calibrated coefficients for the cultivar MGS-BRILHANTE (Table 6) simulated maturity dates and grain yield similarly to the measured data (Fig. 2). The 2-year data used for model evaluation indicate a good model performance to simulate the spring wheat cultivar. The model reproduced the observed maturity dates with an overall R^2 of 0.88, NSE of 0.81, RMSE of 1.52 (days), and d-stat of 0.83. The overall simulated yield had a R^2, NSE, RMSE, and d-stat of 0.85, 0.81, 144.03 (kg ha^{-1}) and 0.77, respectively.

Table 6
Crop genetic parameters for the CSM-CERES-Wheat model.

Cultivar Name	P1V	PID	P5	G1	G2	G3	PHINT
MGS-BRILHANTE	3.46	7.03	509.0	36.1	53.6	1.52	70

3.1.2. Case study 2: GSP calibration of a potato cultivar

The CSM-SUBSTOR-Potato model accurately estimated tuber initiation date, and tuber weight at harvest for the potato cultivar Arizona following calibration (Table 7). Day of tuber initiation was reproduced with an R^2 of 0.96, NSE of 0.78, RMSE of 2.29 (days), and d-stat of 0.75 (Fig. 3). Lastly, tuber dry matter weight was estimated by the model with an R^2 of 1, NSE of 0.98, RMSE of 243.86 (kg ha^{-1}), and d-stat of 0.92.

3.1.3. GSP calibration of a soybean cultivar

The CSM-CROPGRO-Soybean model of DSSAT was used to simulate soybean growth and development. The genetic coefficient for the soybean cultivar S80-P2, was calibrated based on maturity date and grain yield (Table 8). The soybean experiments in Georgia, in 2006, were affected by freezing temperatures, which led to a reduction in crop yield (Bao et al., 2015). For this reason, the grain yield of this given trial was not included, and only the maturity date was utilized for model evaluation as freeze-damaged yield can largely impact the statistical significance of the cultivar calibration results. The maturity date was simulated with an R^2, NSE, RMSE, and d-stat of 0.92, 0.9, 4.37 days and 0.81, respectively (Fig. 4). Crop yield was simulated with an R^2 of 0.75, NSE of 0.72, RMSE of 334.69 kg ha^{-1} and d-stat of 0.72.

3.1.4. Case study 4: Calibration of the cultivar and ecotype coefficients for a maize hybrid

The comparison between the previous genetic coefficients for the DeKalb 52–59 AF2 (VT3) maize hybrid and new calibrations GLUEP showed improvement in the simulation of measured data, particularly when calibrating the genetic coefficients for both the cultivar and ecotype files. Using Anothai et al. (2013) genetic coefficients, the simulations of the anthesis date resulted in an NSE of −0.36, RMSE of 4.95 (days), and d-stat of 0.5. Maturity dates were simulated with a NSE of −0.21, RMSE of 4.95 (days), and d-stat of 0.61. The grain yield had NSE, RMSE, and d-stat values of 0.95, 636 (kg ha^{-1}), and 0.89, respectively (Table 9; Fig. 5).

Upon calibrating the genetic coefficients in the cultivar file with GLUEP, an enhancement in model performance was observed when compared to the genetic coefficients outlined by Anothai et al. (2013). Simulated anthesis dates resulted in an NSE of −1, RMSE of 3.54 (days) and d-stat of 0.5. Correspondingly, maturity dates were simulated with an NSE of −0.75, RMSE of 2.24 (days), and d-stat of 0.78. Moreover, grain yield exhibited further improvement, with NSE, RMSE, and d-stat values of 0.96, 614 (kg ha^{-1}), and 0.89, respectively.

After calibrating the genetic coefficients under the cultivar and ecotype files, the statistical significance of the simulation increased. The

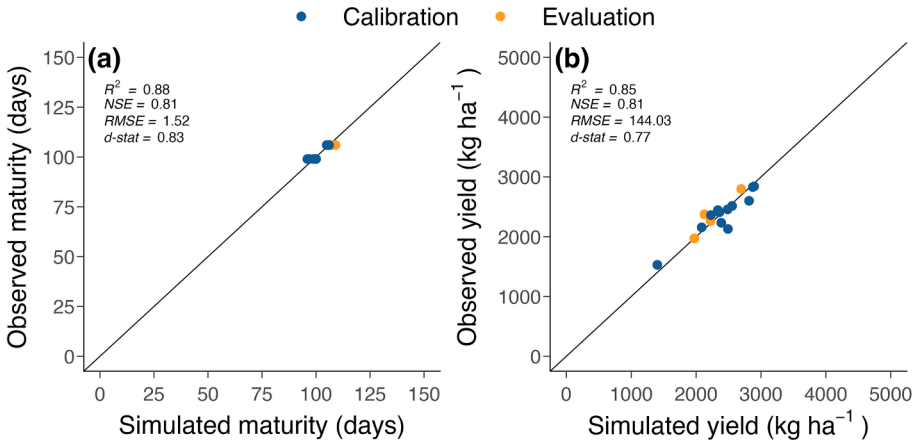


Fig. 2. Identity plot comparing simulated wheat maturity dates **(a)** and grain yield **(b)** from 2013 to 2018 using the CSM-CERES-Wheat model against the measured values from Patos de Minas, Minas Gerais – Brazil. The plot includes the results and statistics for both datasets utilized for calibration (2013, 2014, 2015, and 2017) and evaluation (2016 and 2018) combined. Solid lines represent the 1:1 line.

simulated anthesis dates resulted in an NSE of −0.04, RMSE of 2.55 (days), and d-stat of 0.5. Maturity dates were simulated with an NSE of 0.88, RMSE of 1.58 (days), and d-stat of 0.83. Lastly, grain yield had an NSE, RMSE, and d-stat of 0.96, 614 (kg ha^{-1}), and 0.91, respectively.

3.2. Computational performance and execution time

Genotype calibration utilizing GLUEP increased computation performance and reduced execution time for all computer environments that were analyzed (Figs. 6 and 7). The choice of processor had a significant impact on overall performance. However, the increase in core number resulted in shorter execution time than the single-core approach across all operational systems. Furthermore, GLUEP required substantially less time with a single CPU compared to DSSAT-GLUE, primarily due to the reduction of unnecessary model outputs, which contributed to a faster execution.

Under the Windows operating system, the mean time for exploring the given parameter space using the Intel Core i5-6500 processor was reduced from 1,747 to 153 min (Fig. 6). This represents a 91.2 % reduction in execution time compared to the DSSAT-GLUE. Transitioning to a more robust processor, the Intel Core i7-7700 had twice the maximum number of cores available, and the mean execution time was reduced from 838 to 106 min, a difference of 87.4 %. On the MacOS platform, the Apple M1 Pro chip had a better performance across all CPU configurations. The mean execution time for genetic coefficient calibration using the serial approach was 391 min, while the parallelized approach using ten processors was 36 min, a 91.2 % reduction of the total time required.

The GLUEP also significantly decreased execution time on a high-performance computer cluster (Fig. 7). On average, the time spent conducting the calibration with $N = 50{,}000$ took 1,041 min with a single core using DSSAT-GLUE, while using GLUEP with 120 cores, the calibration took 48 min, a reduction of 95.4 %. With $N = 200{,}000$, a mean time required of 3,425 min with a single core was observed. When using 120 cores, the calibration took 161 min, representing a reduction of 95.3 % in execution time. However, the results show a point of diminishing returns as the number of cores increased beyond which additional cores did not significantly reduce execution time.

The speedup ratio and CPU efficiency were assessed based on the mean execution time for each GLUEP configuration. To isolate the impact of parallelization from source code optimizations, the single-core execution time of GLUEP was used as the baseline in each scenario. This approach allowed for a clear quantification of the parallel processing contribution to the system's overall performance improvements. In the conventional computing environment, Windows 10 systems exhibited limited scalability as the number of cores increased. The maximum speedup ratio was 2.73 when using 4 CPUs with the Intel Core i5-6500 processor and 3.24 with 8 CPUs using the Intel Core i7-7700 processor (Fig. 8). The scalability on the MacOS-based machine was noticeably better with the increase in cores, reaching a maximum of 7.19 with 10 CPUs.

CPU efficiency declined when using more than 2 CPUs for Windows-based systems: −31.8 % CPU efficiency with the Intel Core i5-6500 processor (4 CPUs) and −59.6 % CPU efficiency with the Intel Core i7-7700 processor (8 CPUs). Conversely, the MacOS-based machine had the best performance, with a maximum decrease in CPU efficiency of −28.1 % when using the maximum processing capacity (10 CPUs).

In the high-performance computing environment, the speedup ratio scaled pronouncedly for both RC and HC scenarios as the number of cores utilized for calibration increased (Fig. 9). The speedup ratio reached a value of 12.9 for the RC and 13.6 for the scenario. However, with the increase in the number of cores, the CPU efficiency declined, reaching 10.8 % for the RVCC and 11.3 % for the HVCC scenario when using 120 cores.

Table 7
Crop genetic parameters for the CSM-SUBSTOR-Potato model.

Cultivar Name	G2	G3	PD	P2	TC
Arizona	917.4	21.63	0.524	0.200	15.5

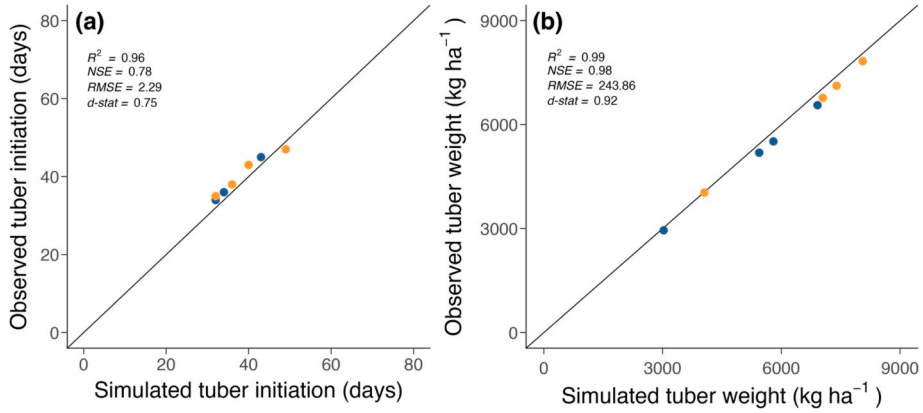


Fig. 3. Identity plot comparing simulated tuber initiation dates **(a)** and tuber dry matter weight **(b)** from two growing seasons (i.e., 2017–18 to 2018–19) using the CSM-SUBSTOR-Potato model against the measured values from Dera Ismail Khan, Khyber Pakhtunkhwa – Pakistan. The plot includes the results and statistics for both datasets utilized for calibration (2017–18) and evaluation (2018–19) combined. Solid lines represent the 1:1 line.

Table 8
Crop genetic parameters for the CSM-CROPGRO-Soybean model.

Cultivar name								
S80-P2								
Genetic coefficients								
CSDL	PPSEN	EM-FL	FL-SH	FL-SD	SD-PM	FL-LF	LFMAX	SLAVR
11.96	0.368	10.11	9.8	15.68	30.91	18	1.068	393.9
SIZLF	XFRT	WTPSD	SFDUR	SDPDV	PODUR	THRSH	SDPRO	SDLIP
201.3	1	0.163	20.85	2.002	10.05	77.31	0.401	0.2

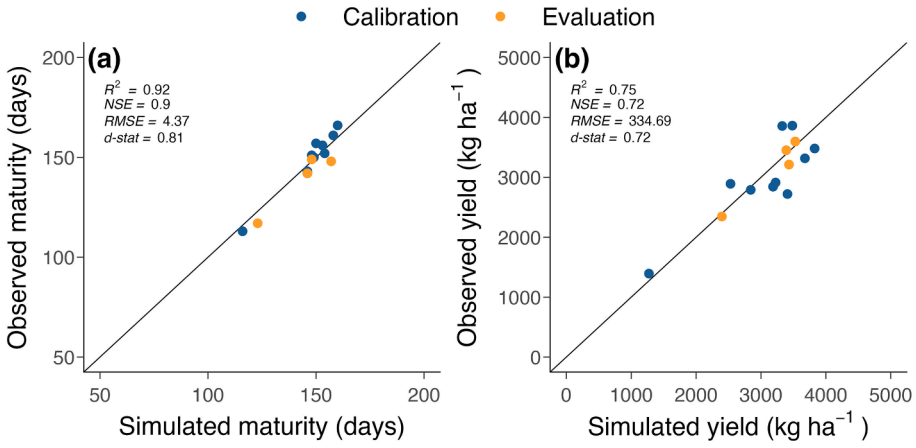


Fig. 4. Identity plot comparing simulated soybean maturity dates **(a)** and grain yield **(b)** from 2005 to 2007 using the CSM-CROPGRO-Soybean model against the measured values from Watkinsville, Williamson, Midville, Plains and Tifton, Georgia, United States. The plot includes the results and statistics for both datasets that were used for calibration (2005 and 2007) and evaluation (2006) combined. Solid lines represent the 1:1 line.

Table 9
Crop genetic parameters for CSM-CERES-Maize model and statistics.

Cultivar Name DeKalb 52–59 AF2				
Genetic coefficients				
	CSM-CERES-Maize genetic parameters	Previously calibrated coefficients (Anothai et al., 2013)	Cultivar calibration	Cultivar and ecotype calibration
Cultivar file coefficients	P1	267.6	260.3	283.6
	P2	0.81	0.509	1.317
	P5	586	598.2	669.2
	G2	960.5	958.9	960.7
	G3	9.42	9.406	16.5
	PHINT	48	65.5	44.55
Ecotype file coefficients	TBASE	8	8	7.4
	TOPT	34	34	33.97
	ROPT	34	34	33.95
	P20	12.5	12.5	12.49
	DJTI	4	4	3.021
	GDDE	6	6	5.28
	DSGFT	170	170	169.9
	RUE	4.2	4.2	3.647
	KCAN	0.85	0.85	0.85
Statistics				
	Goodness-of-fit methods	Previously calibrated coefficients (Anothai et al., 2013)	Cultivar file calibration only	Cultivar and ecotype calibration
Anthesis date	D-Stat	0.5	0.5	0.5
	NSE	−0.36	−1	−0.04
	RMSE	4.95	3.54	2.55
Maturity date	D-Stat	0.61	0.78	0.83
	NSE	−0.21	0.75	0.88
	RMSE	4.95	2.24	1.58
Yield	D-Stat	0.89	0.89	0.91
	NSE	0.95	0.96	0.96
	RMSE	636	614	614

4. Discussion

This study presented the GLUEP software, designed to calibrate the genetic parameters of crop growth models available within the DSSAT-CSM, leveraging the GLUE method. The demand for faster and more robust GSP calibration approaches is essential to simulate the genetic, environmental, and management (G × E × M) interactions with high precision (Wallach et al., 2021b). The GLUEP uses parallel computing to distribute the workload across multiple computer cores and effectively reduce the time required for calibration. Furthermore, GLUEP allows the estimation of model parameters that were unavailable for calibration by previous distributions and flavors of DSSAT-GLUE (He et al., 2010; Hoogenboom et al., 2019; Hyun et al., 2022).

Despite the necessity of establishing prior parameter distribution based on expert knowledge and experimental data, the GLUE method has demonstrated satisfactory parameter estimates even with limited information (Makowski et al., 2002) and across multiple crop growth models (Tremblay and Wallach, 2004; Li et al., 2018; Sheng et al., 2019; Afzal et al., 2024; Dahri et al., 2024). Although the calibration of genetic coefficients using GLUE within the DSSAT-CSM does not consider time series data from field experimentation, the calibration based on end-of-season data can result in a good reproduction of time series observations (Marin et al., 2017). Moreover, another valuable outcome of the GLUE method is the probabilistic distribution of each parameter, which provides insight into the model's sensitivity to individual parameters (Gao et al., 2020).

The GLUEP includes a friendly graphical user interface that resembles the DSSAT-GLUE, eliminating the need for end-users to acquire

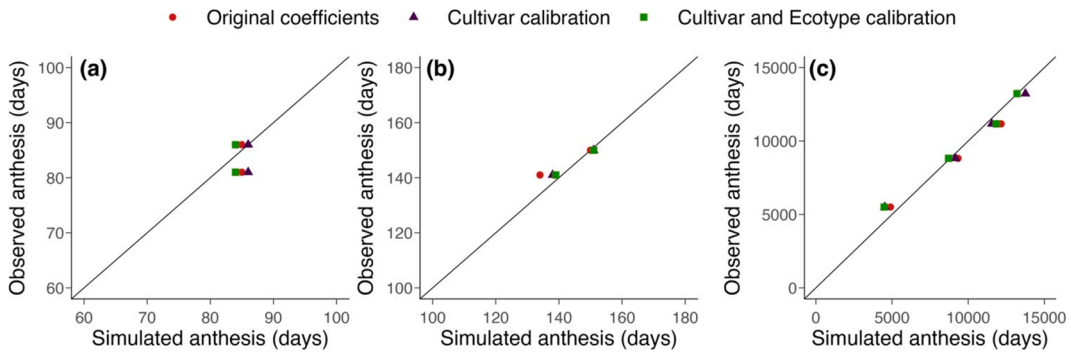


Fig. 5. Identity plot comparing simulated maize anthesis (a) and maturity dates (b) as well as grain yield (c) from 2008 and 2010 using the CSM-CERES-Maize model against the measured values from Greeley, Colorado, United States. The plot includes the results and statistics for both datasets utilized for calibration (2008 under optimal irrigation conditions) and evaluation (2008, suboptimal irrigation conditions, and 2010, optimal and suboptimal irrigation conditions) combined. The plots include the results using the original cultivar coefficients (Anothai et al., 2013), using the cultivar calibration, and using both the cultivar and ecotype calibration. Solid lines represent the 1:1 line.

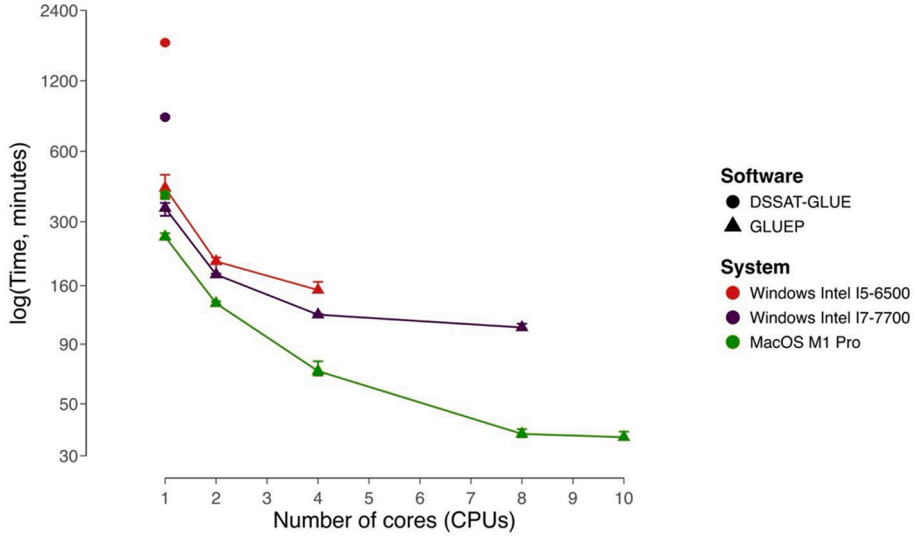


Fig. 6. Mean time required to complete the calibration of genetic-specific parameters (GSP) using DSSAT-GLUE (circle) and GLUEP (triangle) for a conventional computer environment. The simulations were set for a total of 50,000 random parameter sets. Calibration trials were conducted for systems running Windows 10 with Intel Core i5-6500 (red) and Intel Core i7-7700 (purple) processors, and MacOS with M1 Pro chip (green) processor. Error bars represent the range of values obtained for five replications.

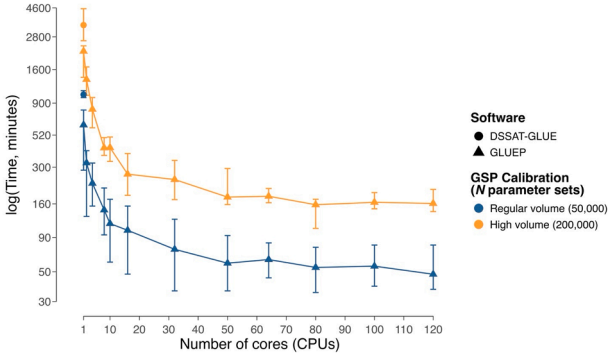


Fig. 7. The mean time required to complete the calibration of genetic-specific parameters (GSP) using DSSAT-GLUE (circle) and GLUEP (triangle) for a high-performance computer environment. The simulations were set for a total of 50,000 (regular volume; blue) and 200,000 (high volume; orange) random parameter sets (N). Error bars represent the range of values obtained for ten replications.

additional knowledge to conduct cultivar calibration using the parallelized system. The system operates on multiple platforms and computer environments, including conventional and high-performance computing setups. Despite the system's broad compatibility, network-based communication via sockets between nodes introduces overhead, mainly due to the continuous data exchange required to manage connections and to ensure data integrity. The overhead is further amplified as objects, functions, and variables must be shared explicitly through communication protocols. The reduction in CPU efficiency with the addition of more CPU cores during calibration suggests a corresponding increase in overhead across all operating systems (Figs. 7 and 8).

One way to mitigate overhead during calibration is to use the forking method (Nyman and Laakso, 2016; R Core Team, 2024), which duplicates the master node environment and shares the entire memory between nodes, thereby minimizing the volume of exchange messages. However, the forking method is incompatible with the Windows system, which would limit GLUEP's portability, particularly given that the DSSAT Shell (Hoogenboom et al., 2019) was designed for Windows. To address this, improving the existing functions for detecting the operational system and applying the most suitable parallel computing method could enhance the performance of the software, especially in high-

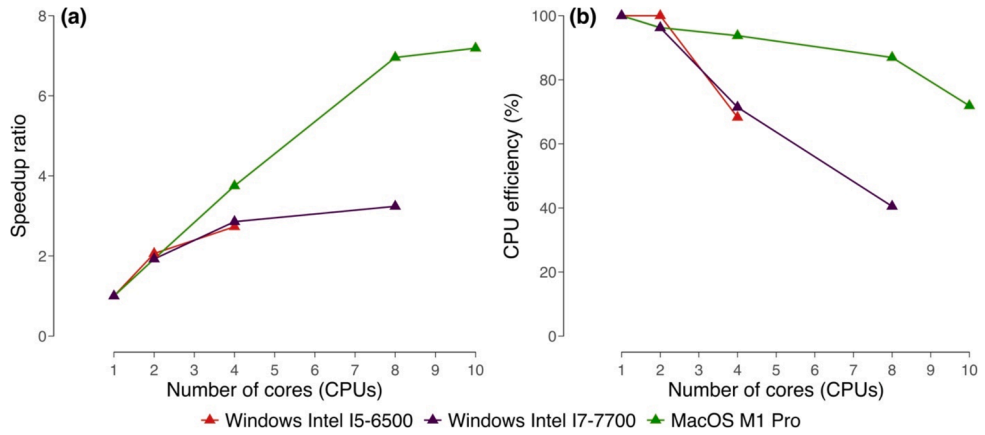


Fig. 8. Average speedup ratio indicating the reduction in execution time by number of cores **(a)** and CPU efficiency **(b)** for convectional computing environments, utilizing the Generalized Likelihood Uncertainty Estimation Parallelized (GLUEP) for genetic-specific parameter calibration. The analysis was performed using computers running Windows 10 with Intel Core i5-6500 (red) and Intel Core i7-7700 (purple) processors, and MacOS with M1 Pro chip (green) processor.

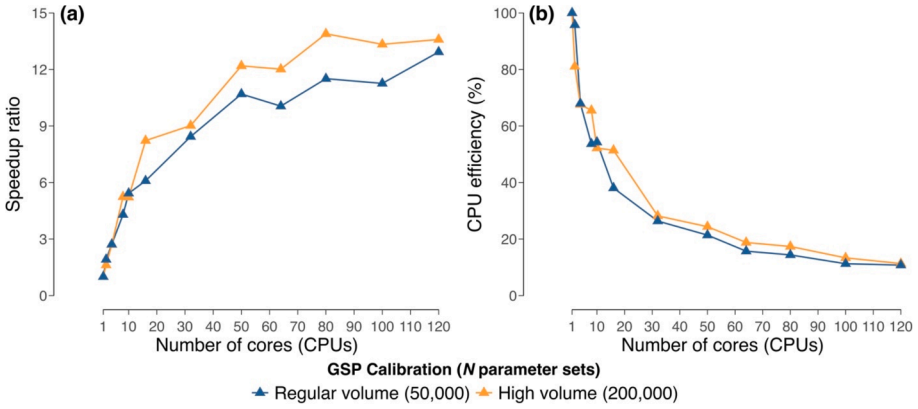


Fig. 9. The speedup ratio indicating the reduction in execution time by the number of cores **(a)** and CPU efficiency **(b)** for a high-performance computing environment, utilizing the Generalized Likelihood Uncertainty Estimation Parallelized (GLUEP) for genetic-specific parameter calibration. The analysis was performed using *N* parameter sets of 50,000 (regular volume; blue) and 200,000 (high volume; orange).

performance computing clusters running on Unix-like operational systems.

Four case studies using different crops and models showcase the versatility and practical applications of GLUEP for calibrating crop genetic parameters. With the continuous expansion and addition of crop models to represent different crop genotypes within the DSSAT-CSM, the GLUEP flexibility facilitates the integration of new models for parameter calibration. One case study on the DeKalb 52–59 AF2 (VT3) maize hybrid exemplified the potential use of calibrating the genetic coefficients defined within the ecotype file. Model users are recommended not to modify ecotype files as the genetic coefficients are often shared among similar genotypes, and the manipulation of these files is reserved for model developers (Hoogenboom et al., 2019). However, GLUEP integration with the ecotype files enables any user to seamlessly calibrate the ecotype-related coefficients of the crop models within DSSAT-CSM.

In each case study, the value of *N*, the number of parameter sets generated, varied to showcase the software's capabilities in calibrating genetic coefficients. Exploring the parameter space for the optimal parameter set depends on the magnitude of *N*, the number of genetic coefficients being calibrated for a given model, and the ranges associated with each coefficient. Therefore, defining an optimal value of *N* for obtaining statistically significant GSP calibration results is highly uncertain. However, it is possible to infer that increasing the size of *N* broadens the parameter space exploration, enhancing the likelihood of converging to the global optimum parameter set.

The software performance analysis was conducted in multiple systems employing different core configurations for a series of calibration procedures, incorporating replicates to ensure result reliability and variability assessment. Mean values were used to indicate any performance improvement. Across all platforms and systems examined, GLUEP decreased the time required to conduct GSP calibration. Such performance enhancements can further enable and accelerate studies utilizing crop growth models for crop phenotyping and prediction of G × E × M interactions concerning new cultivar and accessions (Jones et al., 2011; Bustos-Korts et al., 2019; Wallach et al., 2021b).

The mean execution time peaked when utilizing a single core, as GLUEP behaves as a serial program, executing the calibration steps sequentially. When comparing GLUEP's single-core results to DSSAT-GLUE, the wall clock time was reduced by 33.8 % to 76.1 % in conventional computing environments and by 36.1 % to 40.5 % in a high-performance computing environment. This improvement is attributed to code optimization, which suppressed unnecessary crop model outputs during GSP calibration. The distributed system reduced the wall clock time as cores increased. A more pronounced reduction was observed with two to four cores in the conventional computer environments analyzed. In the HPC cluster, a significant decrease in execution time was observed between two to twenty cores. However, the average time reduction after twenty cores decreased and was nearly identical to the 120 cores configuration. The same trend was present in the conventional computer environments after four cores for the Macbook M1 Pro and Windows 10 Intel Core i7-7700. This is expected by any parallel

algorithm, where performance gains and CPU efficiency diminishes as core count increases due to overhead and additional communication between CPUs. Furthermore, R's single-threaded internal structure can lead to an underutilization of computational resources, as modern CPUs often support multithreading for executing multiple threads concurrently on each core. While there are approaches to enable multithreading with R, they require installing and using external programs that replace some of the R internal libraries. However, these significant modifications demand careful consideration, as they can disrupt the compatibility with R packages that can result unexpected system response and, thus, increase the complexity of system maintenance.

The CPU efficiency and speedup ratio metrics indicated that the Macbook M1 Pro had the best performance under the performance analysis for conventional computing environments. However, the Windows machines used in this study had low-end CPUs compared to the MacOS system, influencing the calibration performance. Furthermore, Thorp et al. (2012) highlighted that Windows-based systems have slower performance when running crop growth models, which can further affect the overall performance. For HPC environments, the CPU efficiency is likely to increase with a high volume of model runs for GSP calibration as overhead becomes smaller relative to the algorithm execution. This can be inferred based on the higher speedup ratio in the high-volume GSP calibration compared to the regular-volume GSP calibration.

The number of parameters in the parameter space is also expected to impact calibration time. The significant variation in the HPC can be attributed to the number of tasks the node was simultaneously executing. Lastly, the posterior distribution analysis is a serial task as each individual processor converges to conduct the analysis. Therefore, the value of N can also affect the time required for the algorithm to merge the outputs and identify the parameter set with the highest probability value.

5. Conclusions

Calibrating genetic-related coefficients is a fundamental step for ensuring accurate predictions of crop growth models. In this study, we demonstrated the efficacy of the Generalized Likelihood Uncertainty Estimation Parallelized (GLUEP) in streamlining the calibration process, significantly decreasing execution time and enhancements in crop model parameter estimation. The tool was evaluated using four case studies centered on different crops and crop growth models within the DSSAT Crop Modeling Ecosystem, demonstrating the GLUEP capability of calibrating multiple genotypes across different crop growth models. The deployment in low- and high-end computing systems using different operational systems indicates the compatibility and adaptability of the GLUEP software. Ultimately, GLUEP can potentially improve the calibration of GSPs and provide a more accurate and reliable representation of crop growth and development through crop models. Future research includes enhancing the GLUEP structure to detect the operational system and applying the forking method for parallel processing, which could further improve the system efficiency on Unix-like systems. Additionally, the use of multithreading should be explored to assess its effect on computational performance on GLUEP during calibration of GSPs.

CRediT authorship contribution statement

Thiago Berton Ferreira: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Vakhtang Shelia:** Writing – review & editing, Validation, Software. **Cheryl Porter:** Writing – review & editing, Validation. **Patricia Moreno Cadena:** Validation, Software. **Montserrat Salmeron Cortasa:** Writing – review & editing, Validation. **Muhammad Sohail Khan:** Writing – review & editing, Resources, Data curation. **Willingthon Pavan:** Validation, Software. **Gerrit Hoogenboom:** Writing – review & editing, Validation, Supervision, Resources, Methodology.

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.

References

Afzal, O., Ahmed, M., Shabbir, G., Ahmed, S., Hoogenboom, G., 2024. CSM-CROPGRO model to simulate safflower phenological development and yield. Int. J. Biometeorol. 1–16.

Ahmed, M., Ahmad, S., Raza, M.A., Kumar, U., Ansar, M., Shah, G.A., Parsons, D., Hoogenboom, G., Palosuo, T., Seidel, S., 2020. Models calibration and evaluation. Systems Modeling 151–178.

Akhavizadegan, F., Ansarifar, J., Wang, L., Huber, I., Archontoulis, S.V., 2021. A time-dependent parameter estimation framework for crop modeling. Sci. Rep. 11 (1), 11437.

Alderman, P.D., Stanfill, B., 2017. Quantifying model-structure-and parameter-driven uncertainties in spring wheat phenology prediction with Bayesian analysis. Eur. J. Agron. 88, 1–9.

Anothai, J., Patanothai, A., Jogloy, S., Pannangpetch, K., Boote, K.J., Hoogenboom, G., 2008. A sequential approach for determining the cultivar coefficients of peanut lines using end-of-season data of crop performance trials. Field Crop Res 108, 169–178.

Anothai, J., Soler, C.M.T., Green, A., Trout, T.J., Hoogenboom, G., 2013. Evaluation of two evapotranspiration approaches simulated with the CSM–CERES–Maize model under different irrigation strategies and the impact on maize growth, development and soil moisture content for semi-arid conditions. Agric. For. Meteorol. 176, 64–76.

Ascari, J.P., Barro, J.P., Santana, F.M., Padua, J.M., Maciel, J.L., Lau, D., Del Ponte, E.M., 2021. Sequential post-heading applications for controlling wheat blast: A 9-year summary of Fungicide performance in Brazil. Plant Dis. 105 (12), 4051–4059.

Bao, Y., Hoogenboom, G., McClendon, R.W., Paz, J.O., 2015. Potential adaptation strategies for rainfed soybean production in the south-eastern USA under climate change based on the CSM-CROPGRO-Soybean model. J. Agric. Sci. 153 (5), 798–824.

Bennett, J.M., Mutti, L.S.M., Rao, P.S.C., Jones, J.W., 1989. Interactive effects of nitrogen and water stresses on biomass accumulation, nitrogen uptake, and seed yield of maize. Field Crop Res 19 (4), 297–311.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6 (3), 279–298.

Beven, K., Binley, A., 2014. GLUE: 20 years on. Hydrol. Process. 28 (24), 5897–5918.

Boote, K.J., Kropff, M.J., Bindraban, P.S., 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. Agr. Syst. 70 (2–3), 395–420.

Boote, K.J., Jones, J.W., Hoogenboom, G., 2021. Incorporating realistic trait physiology into crop growth models to support genetic improvement. Silico Plants 3 (1), diab002.

Buddhaboon, C., Jintrawet, A., Hoogenboom, G., 2018. Methodology to estimate rice genetic coefficients for the CSM-CERES-Rice model using GENCALC and GLUE genetic coefficient estimators. J. Agric. Sci. 156 (4), 482–492.

Bustos-Korts, D., Boer, M.P., Malosetti, M., Chapman, S., Chenu, K., Zheng, B., Van Eeuwijk, F.A., 2019. Combining crop growth modeling and statistical genetic modeling to evaluate phenotyping strategies. Front. Plant Sci. 10, 481587.

Dahri, S.H., Shaikh, I.A., Talpur, M.A., Mangrio, M.A., Dahri, Z.H., Hoogenboom, G., Knox, J.W., 2024. Modelling the impacts of climate change on the sustainability of rainfed and irrigated maize in Pakistan. Agric Water Manag 296, 108794.

Doherty, J.E., Hunt, R.J., Tonkin, M.J., 2010, Approaches to highly parameterized inversion: A guide to using PEST for model-parameter and predictive-uncertainty analysis: U.S. Geological Survey Scientific Investigations Report 2010–5211, pp. 71.

Gao, Y., Wallach, D., Liu, B., Dingkuhn, M., Boote, K.J., Singh, U., Asseng, S., Kahveci, T., He, J., Zhang, R., Confalonieri, R., Hoogenboom, G., 2020. Comparison of three calibration methods for modeling rice phenology. Agric. For. Meteorol. 280, 107785.

GLUEP Source code. <https://github.com/DSSAT/GLUE> (Accessed on 06 September 2024).

Grama, A.Y., Gupta, A., Kumar, V., 1993. Isoefficiency: Measuring the scalability of parallel algorithms and architectures. IEEE Parallel Distribut. Technol. Syst. Appl. 1 (3), 12–21.

Harrison, M.T., Roggero, P.P., Zavattaro, L., 2019. Simple, efficient and robust techniques for automatic multi-objective function parameterisation: Case studies of local and global optimisation using APSIM. Environ. Model. Softw. 117, 109–133.

He, J., Jones, J.W., Graham, W.D., Dukes, M.D., 2010. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method. Agr. Syst. 103 (5), 256–264.

Hoogenboom, G., Porter, C.H., Boote, K.J., Shelia, V., Wilkens, P.W., Singh, U., White, J. W., Asseng, S., Lizaso, J.I., Moreno, L.P., Pavan, W., Ogoshi, R., Hunt, L.A., Tsuji, G. Y., Jones, J.W., 2019. The DSSAT crop modeling ecosystem. In Advances in crop modelling for a sustainable agriculture. In: Advances in crop modelling for a sustainable agriculture. Burleigh Dodds Science Publishing, pp. 173–216.

Hoogenboom, G., Porter, C.H., Shelia, V., Boote, K.J., Singh, U., Pavan, W., Oliveira, F.A. A., Moreno-Cadena, L.P., Ferreira, T.B., White, J.W., Lizaso, J.I., Pequeno, D.N.L., Kimball, B.A., Alderman, P.D., Thorp, K.R., Cuadra, S.V., Vianna, M.S., Villalobos, F. J., Batchelor, W.D., Asseng, S., Jones, M.R., Hopf, A., Dias, H.B., Hunt, L.A., Jones, J. W., 2023. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.8.2. DSSAT Foundation, Gainesville, Florida, USA.

Hoogenboom, G., Porter, C.H., Shelia, V., Boote, K.J., Singh, U., Pavan, W., Oliveira, F.A. A., Moreno-Cadena, L.P., Ferreira, T.B., White, J.W., Lizaso, J.I., Pequeno, D.N.L., Kimball, B.A., Alderman, P.D., Thorp, K.R., Cuadra, S.V., Vianna, M.S., Villalobos, F. J., Batchelor, W.D., Asseng, S., Jones, M.R., Hopf, A., Dias, H.B., Jintrawet, A., Jaikla, R., Memic, E., Hunt, L.A., Jones, J.W., 2024. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.8.5 (www.DSSAT.net). DSSAT Foundation, Gainesville, Florida, USA.

Hunt, L. A., Boote, K. J., 1998. Data for model operation, calibration, and evaluation. Understanding options for agricultural production, 9-39.

Hunt, L.A., Pararajasingham, S., Jones, J.W., Hoogenboom, G., Imamura, D.T., Ogoshi, R. M., 1993. GENCALC - Software to facilitate the use of crop models for analyzing field experiments. Agronomy Journal 85 (5), 1090–1094.

Hyun, S., Park, J.Y., Kim, J., Fleisher, D.H., Kim, K.S., 2022. GLUEOS: A high performance computing system based on the orchestration of containers for the GLUE parameter calibration of a crop growth model. Comput. Electron. Agric. 197, 106906.

Jha, P.K., Ines, A.V., Han, E., Cruz, R., Prasad, P.V., 2022. A comparison of multiple calibration and ensembling methods for estimating genetic coefficients of CERES-Rice to simulate phenology and yields. Field Crop Res. 284, 108560.

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.

Jones, J.W., He, J., Boote, K.J., Wilkens, P., Porter, C.H., Hu, Z., 2011. Estimating DSSAT cropping system cultivar-specific parameters using Bayesian techniques. Method. Introduc. Syst. Model. Agricult. Res. 2, 365–393.

Khan, M.S., Hoogenboom, G., Gillani, S.M., Shah, A. S., Khan, I., 2024. Effects of planting date and genotypes on potato growth and yield determination in a sub-tropical continental growing environment. Accepted at Potato Research.

Li, Y., Gholami, H., Song, Y., Fathabadi, A., Malakooti, H., Collins, A.L., 2020. Source fingerprinting loess deposits in Central Asia using elemental geochemistry with Bayesian and GLUE models. Catena 194, 104808.

Li, Z., He, J., Xu, X., Jin, X., Huang, W., Clark, B., Yang, G., Li, Z., 2018. Estimating genetic parameters of DSSAT-CERES model with the GLUE method for winter wheat (*Triticum aestivum* L.) production. Comput. Electron. Agric. 154, 213–221.

Ma, H., Malone, R.W., Jiang, T., Yao, N., Chen, S., Song, L., Feng, H., Yu, Q., He, J., 2020. Estimating crop genetic parameters for DSSAT with modified PEST software. Eur. J. Agron. 115, 126017.

Makowski, D., Wallach, D., Tremblay, M., 2002. Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods. Agronomie 22 (2), 191–203.

Marin, F., Jones, J.W., Boote, K.J., 2017. A stochastic method for crop models: including uncertainty in a sugarcane model. Agron. J. 109 (2), 483–495.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J. W., Rötter, R. P., Boote K. J., Ruane, A. C., Horburn P. J., Cammarano, D., Hatfield, J. L., Rosenzweig, C., Aggarwal, P. K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A. J., Doltra J., Gayler, S., Goldberg, R., Grant, R. F., Heng, L., Hooker, J., Hunt, L. A., Ingwesen J., Izaurralde, R. C., Kersebaum, K., C., Müller, C., Kumar, S. N., Nendel C., O'leary, G., Olesen, J. E., Osborne, T. M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M. A., Shcherbak, I., Steduto, P., Stöckle, C. O., Stratonovitch, P., Streck, T., Supit, I., Tao., F., Travasso, M., Waha, K., White, J. W., Wolf, J., 2015. Multimodel ensembles of wheat growth: many models are better than one. Global Chang. Biol., 21(2), 911-925.

Memic, E., Graeff, S., Boote, K.J., Hensel, O., Hoogenboom, G., 2021. Cultivar coefficient estimator for the cropping system model based on time-series data: a case study for soybean. Trans. ASABE 64 (4), 1391–1402.

Mirzaei, M., Huang, Y.F., El-Shafie, A., Shatirah, A., 2015. Application of the generalized likelihood uncertainty estimation (GLUE) approach for assessing uncertainty in hydrological models: a review. Stoch. Env. Res. Risk A. 29, 1265–1273.

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.

Nyman, L., Laakso, M., 2016. Notes on the history of fork and join. IEEE Ann. Hist. Comput. 38 (3), 84–87.

Perkins, H. F., Hook, J. E., Barbour, N. W., 1986. Soil characteristics of selected areas of the Coastal Plains Experiment Station and ABAC research farms.

R Core Team. (2024). Package 'parallel'. R Foundation for Statistical Computing. https://www.r-project.org/.

Ragab, R., Kaelin, A., Afzal, M., Panagea, I., 2020. Application of Generalized Likelihood Uncertainty Estimation (GLUE) at different temporal scales to reduce the uncertainty level in modelled river flows. Hydrol. Sci. J. 65 (11), 1856–1871.

Ramalakshmi, E., Kompala, N. (2017). Multi-threading image processing in single-core and multi-core CPU using R language. In 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT) (pp. 1-5). IEEE.

Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S.V., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci. Rep. 11 (1), 1606.

Sheng, M., Liu, J., Zhu, A.X., Rossiter, D.G., Liu, H., Liu, Z., Zhu, L., 2019. Comparison of GLUE and DREAM for the estimation of cultivar parameters in the APSIM-maize model. Agric. For. Meteorol. 278, 107659.

Thorp, K.R., White, J.W., Porter, C.H., Hoogenboom, G., Nearing, G.S., French, A.N., 2012. Methodology to evaluate the performance of simulation models for alternative compiler and operating system configurations. Comput. Electron. Agric. 81, 62–71.

Tremblay, M., Wallach, D., 2004. Comparison of parameter estimation methods for crop models. Agronomie 24 (6–7), 351–365.

Wallach, D., Palosuo, T., Thorburn, P., Hochman, Z., Gourdain, E., Andrianasolo, F., Asseng, S., Basso, B., Buis, S., Crout, N., Dibari, C., Dumont, B., Ferrise R., Gaiser, T., Garcia, C., Gayler S., Ghahramani A., Hiremath, S., Hoek, S., Horan, H, Hoogenboom, G., Huang, M., Jabolun M., Jansson, P., Jing, Q., Justes, E., Kersebaum, K. C., K. C., Klosterhalfen A., Launay, M., Lewan, E., Luo, Q., Maestrini, B., Mielenz, H., Moriondo, M., Zadeh, Padovan, G., Olesen, J. E., Poyda, A., Priesack, E., Pullens, J. W. M., Qian, B., Shütze, N., Shelia, V., Souissi A., Specka, X., Srivastava, A. K., Stella, T., Streck T., Trombi, G., Wallor, E., Wang, J., Weber, T. K. D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao C., Zhu, Y., Seidel, S. J., 2021a. The chaos in calibrating crop models: Lessons learned from a multi-model calibration exercise. Environmental Modelling and Software, 145, 105206.

Wallach, D., Palosuo, T., Thorburn, P., Gourdain, E., Asseng, S., Basso, B., Buis, S., Neil, C., Dibari, C., Dumont, B., Ferrise, R., Gaiser, T., Garcia, C., Gayler S., Ghahramani A., Hochman Z., Hoek, S., Hoogenboom, G., Horan, H., Huang, M., Jabolun M., Jing, Q., Justs, E., Kersebaum, K. C., Klosterhalfen A., Launay, M., Luo Q., Maestrini, B., Mielenz, H., Moriondo, M., Zadeh, H. N., Olesen J., E., Poyda, A., Priesack, E., Pullens, J. W. M., Qian, B., Shütze, N., Shelia, V., Souissi A., Specka, X., Srivastava, A. K., Stella, T., Streck T., Trombi, G., Wallor, E., Wang, J., Weber, T. K. D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao C., Zhu, Y., Seidel, S. J., 2021b. How well do crop modeling groups predict wheat phenology, given calibration data from the target population? Europ. J. Agron., 124, 126195.

Wallach, D., Palosuo, T., Thorburn, P., Mielenz, H., Buis, S., Hochman, Z., Seidel, S.J., 2023. Proposal and extensive test of a calibration protocol for crop phenology models. Agron. Sustain. Dev. 43 (4), 46.

White, J.W., Hoogenboom, G., 1996. Simulating effects of genes for physiological traits in a process-oriented crop model. Agron. J. 88 (3), 416–422.

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res. Oceans 90 (C5), 8995–9005.

Xavier, A.C., Scanlon, B.R., King, C.W., Alves, A.I., 2022. New improved Brazilian daily weather gridded data (1961–2020). Int. J. Climatol. 42 (16), 8390–8404.