https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

LOGO: A hybrid approach for calibration of crop model parameters using both local and global methods in tandem

Shinwoo Hyun[1], Junhwan Kim[2], Young Sang Joh[3], Kwang Soo Kim[1,3], Gerrit Hoogenboom[4,5]

[1] Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, 08826, Republic of Korea
[2] Korea National University of Agriculture and Fisheries, Jeonju-si, Jeollabuk-do, 54874, Republic of Korea
[3] Department of Agriculture, Forestry, and Bioresources, Seoul National University, Seoul, 08826, Republic of Korea
[4] Agricultural and Biological Engineering Department, University of Florida, Gainesville, FL, 32611-0570, USA
[5] Global Food Systems Institute, University of Florida, Gainesville, FL, 32611-0910, USA

Correspondence to: Kwang Soo Kim (luxkwang@snu.ac.kr)

Abstract. Calibration of cultivar parameters is essential in ensuring the reliability of process-based crop models. Although local optimization methods have been widely used for parameter calibration, they can be sensitive to initial parameter values. Here we propose Local Optimization Global Oriented (LOGO) framework, which integrates a local search algorithm with posterior distribution assessment derived from the global search algorithm. In the present study, we examined the effects of initial values and site diversity on calibration outcomes with pseudo-observation datasets generated across six sites representing rice production environments in Asia. Our results indicated that estimated phenology-related parameters were relatively robust, whereas yield-related parameters had greater variability depending on the choice of initial values. Application of the LOGO framework reduced the distance between true and estimated parameter values. These findings provide a systematic basis for developing more robust calibration protocols and highlight the importance of integrating multiple optimization strategies.

1. Introduction

Crop growth models have been used to simulate plant development, biomass accumulation, and yield formation under diverse environmental conditions (Timlin et al., 2024; Archontoulis et al., 2020). These models serve as essential tools for evaluating crop management performance, assessing climate change impacts, and supporting agronomic decision-making (Yoo et al., 2025; Rezaei et al., 2023). They also facilitate analysis related to policy designs on food security (Gavasso-Rita et al., 2024). The reliability of crop models is, therefore, critical for their effective application in informing agronomic management, guiding policy decisions, and advancing climate adaptation strategies.

The accuracy of crop growth simulations is dependent on, in part, the reliability of parameter estimates derived through calibration (Hyun et al., 2022). Calibration involves optimizing model parameters to maximize the degree of agreement between observed and simulated outputs across a range of environmental and management conditions (Gao et al., 2021; Potgieter et al., 2021). A number of calibration protocols have been suggested to promote efficiency and reliability in

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

parameter estimation (Archontoulis et al., 2014; Gao et al., 2020). In particular, Seidel et al. (2018) proposed guidelines for calibration practices, which can improve reliability of crop growth simulations in a standard way. Wallach et al. (2024) further outlined a protocol for crop model calibration, including step-wise selection of parameters and evaluation against independent datasets.

Local optimization methods are commonly used in cultivar parameter calibration due to their computational efficiency and ease of implementation (Wallach et al., 2021). Typically, this process entails adjusting parameters related to phenological and growth traits to match observed crop responses (Wallach et al., 2021). For example, GENCALC, which iteratively adjusts parameter values to minimize errors in parameter-related traits (Hunt et al., 1993), has been used to estimate genetic coefficients in the DSSAT model (Buddhaboon et al., 2018). Similarly, the Nelder-Mead algorithm has been used for the Agricultural Model Intercomparison and Improvement Project (AgMIP) (Wallach et al., 2025). However, local optimization can be sensitive to the choice of initial parameter values. If suboptimal initial parameter values are selected, the optimization process may converge to local rather than global minima, thereby compromising the robustness and portability of the calibration results (Jeraj et al., 2003).

Global calibration techniques such as the Generalized Likelihood Uncertainty Estimation (GLUE) offer an alternative, as they explore the full parameter space without requiring initial estimates near the global optima (Shoarinezhad et al., 2020; Ferreira et al., 2024). However, GLUE does not include an iterative fitting procedure, making it challenging to refine parameter estimates toward an optimal solution. The Metropolis–Hastings (M-H) algorithm provides probabilistic frameworks for exploring parameter distributions. While effective in estimating parameter uncertainty, this method is computationally intensive and not easily parallelized, limiting their scalability on modern high-performance computing architectures (Jacob et al., 2011).

Here, we propose the hybrid calibration strategy that integrates local and global optimization approaches, referred to as Local Optimization Global Oriented (LOGO) framework, to improve the robustness, accuracy, and portability of parameter estimation for crop growth models. The objectives of this study were to address key limitations associated with traditional calibration methods. To this end, we investigated three research questions: (1) How do initial parameter values influence the robustness of cultivar parameter calibration using the local optimization method? (2) How does the availability of observation data across multiple sites affect the reliability and transferability of estimated parameters? (3) Can the use of prior and posterior distributions derived from the combination of sites enhance the reliability of parameter estimates through the GLUE method? The answers to these questions would provide insights into a cohesive parameter calibration framework that balances computational efficiency with accuracy. This would ultimately improve the predictive performance of process-based crop models under diverse environmental and management conditions. Such a framework can enhance the utility of process-based crop models for decision-making in crop production systems and agricultural policy development.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

2. Materials and Methods

2.1. Pseudo-observation data

The experimental sites were selected to represent diverse climate conditions under which rice is grown across Asia (Table 1). Los Baños (Philippines), Ludhiana (India), and Nanjing (China) were designated as core localities, which was used in Li et al. (2015). Gazipur (Bangladesh), Milyang (Korea) and Suwon (Korea) were included to enhance geographic and environmental diversity during the calibration processes. These locations had a wide range of climate zones (see Supplementary Information 1), which was identified using the Köppen climate classification system (Beck et al., 2023).

Pseudo-observation datasets that emulate real field observations were generated at the experimental sites for the given planting dates (Fig. 1). These surrogates of actual observations allowed for the evaluation of the uncertainties caused only by the calibration procedures. The crop model of interest was operated to prepare those data at each site of interest. The datasets included phenology (e.g., panicle initiation, heading date, and maturity date) and growth (e.g., biomass, grain number, and yield) data. In particular, biomass data were retrieved from the outputs of the model at the first date of mid tillering, panicle initiation, flowering, and grain filling stages, respectively, under the assumption that its measurements were available at field experiments (Peng et al., 2000). The simulation periods were divided into two decades. Data from 1991–2000 were used for calibration, and those from 2001–2010 were reserved for evaluation of the calibrated parameters.

Table 1. Sites where parameter calibration experiments were conducted.

Location	Latitude	Longitude	Climate type	Planting date	Transplanting date	Reference
Los Banos, Philippines	14° 6’ N	121° 9’ E	Tropical, monsoon (Am)[a]	01 January	15 January	Peng et al., 2000
Gazipur, Bangladesh	23° 59’ N	90° 24’ E	Temperate, dry winter, hot summer (Cwa)	30 June	30 July	Rahman et al., 2019
Ludhiana, India	30° 54’ N	75° 48’ E	Arid, steppe, hot (BSh)	01 June	25 June	Kaur et al., 2023
Nanjing, China	32° 56’ N	118° 59’ E	Temperate, hot summer, no dry season (Cfa)	7 June	28 June	Dou et al., 2021
Milyang, Korea	35° 29’ N	128° 44’ E	Temperate, dry winter, hot summer (Cwa)	10 May	05 June	Kim et al., 2020
Suwon, Korea	37° 15’ N	126° 59’ E	Cold, dry winter, hot summer (Dwa)	20 April	20 May	Kim et al., 2020

[a] Alphabet code within parenthesis indicates Köppen climate classification.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

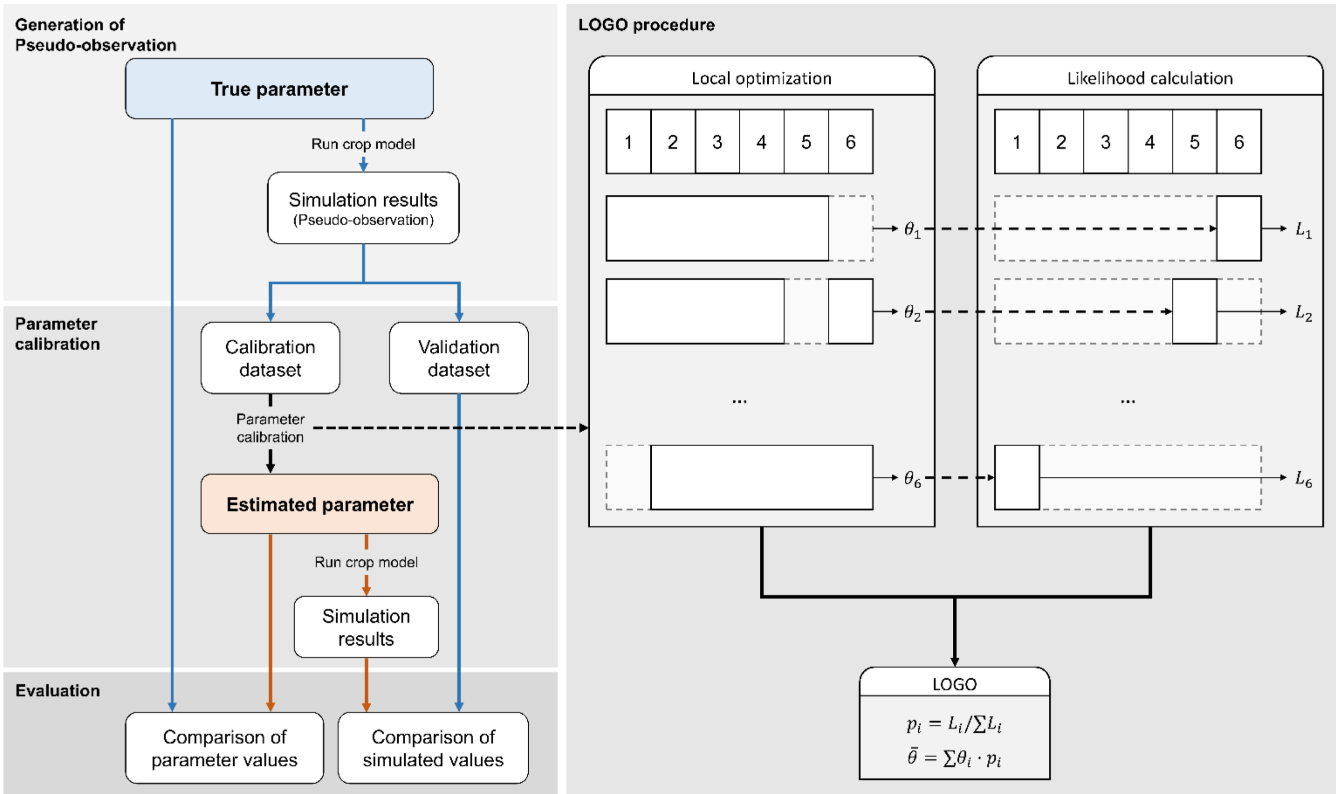

Figure 1: Flow chart for parameter comparison between true and estimated parameters. A set of true parameters was first defined and used to generate pseudo-observation data, from which the parameters were then estimated through the calibration procedures. Within the Local Optimization and Global Oriented (LOGO) framework, the local optimization was repeated over combinations of site subsets to determine the parameter values, $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$, …,$\boldsymbol{\theta_n}$. These parameters constitute the parameter search space for the global search method where likelihood values are determined to derive the estimates of parameter.

2.2. Crop growth simulations

All simulations were carried out using the Cropping System Model (CSM)-CERES-Rice model, part of the Decision Support System for Agrotechnology Transfer (DSSAT v4.8.2; Hoogenboom et al., 2023). This process-based model integrates physiological processes such as phenology, photosynthesis, nutrient uptake, and water balance to simulate daily crop-soil-atmosphere interactions (Hoogenboom et al., 2019). The CSM-CERES-Rice model was used for the generation of pseudo-observation data and calibration of cultivar parameters in the present study.

The cultivar IR72 was selected for the crop growth simulations. IR72 has been widely adopted due to its high yield potential, tolerance to diseases and pests, and adaptability to a variety of agro-climatic conditions (Zhu et al., 2002). This cultivar served as an ideal baseline for the assessment of parameter calibration because it has been used for crop model development (Peng et al., 1999). It is also included in the standard cultivar library for the CSM-CERES-Rice model. The default parameters for IR72 were adopted as the true parameter values, which was denoted by $\boldsymbol{P}_o$.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

Crop management practices were set using existing experiments and recommendations (Table 1). The site-specific planting dates were determined based on previous studies. It was assumed that nitrogen fertilizer application followed a three-stage application protocol as recommended by the International Rice Research Institute (IRRI) (http://www.knowledgebank.irri.org/ericeproduction/pop_up_Nitrogen_application.htm). Fertilizer applications were timed to coincide with critical growth stages. For example, basal fertilizer was applied three days prior to transplanting at the rate of 45 kg N ha^{-1}. Fertilizers at tillering and panicle initiation were applied 10 and 50 days after transplanting at the rate of 25 and 35 kg N ha^{-1}, respectively. Irrigation was assumed to be optimal throughout the growing season because rice is commonly grown in lowland areas.

Weather input datasets for the crop growth simulations were prepared using the AgERA5 global reanalysis data, which is designed to support agricultural modeling studies (Boogaard et al., 2020). AgERA5 provides high-resolution meteorological data at an approximate resolution of 10 km at a global scale, offering continuous and spatially consistent weather data. Daily maximum and minimum temperatures, precipitation, solar radiation, and wind speed were extracted for the study sites for a 30-year period from 1991 to 2010. These data were obtained from the Copernicus Climate Data Store (CDS; https://cds.climate.copernicus.eu/datasets/sis-agrometeorological-indicators, accessed on 05 March 2025).

Soil input data were prepared using the SoilGrids database (https://soilgrids.org), which provides high-resolution global soil profile data to a depth of 2 m with a spatial resolution of 250 m (Hengl et al., 2017). This dataset includes soil properties such as texture, bulk density, and hydraulic characteristics, which are crucial for reliable simulation of soil-plant interactions. In the present study, these soil attributes were processed to generate soil input files for the CSM-CERES-Rice model using algorithms suggested by Kim et al. (2018).

2.3. Local optimization procedures

Parameter calibration experiments were conducted to estimate the value of $\boldsymbol{P_o}$. At first, the CroptimizR package was used to solve the inverse problem. CroptimizR is an R software package designed to facilitate the calibration of process-based crop models, which has been used in the AgMIP Calibration project (Buis et al., 2024). By default, it is based on the Nelder-Mead method, which iteratively adjusts parameter values (Nelder and Mead, 1965). This approach minimizes the discrepancies between simulated and observed data using Akaike Information Criterion (AIC).

A wrapper for the CSM-CERES-Rice model was implemented to couple the model with CroptimizR. CroptimizR provides a flexible framework that can interface seamlessly with any crop model through custom R wrappers (Buis et al., 2024). This wrapper standardizes the process of running simulations by exchanging inputs and outputs in a consistent format. This allows for users to force specific parameter values into their crop model simulations and retrieve the resulting outputs in a predictable structure. This standardized interface not only simplifies the integration of diverse models into the calibration workflow but also ensures subsequent analyses and diagnostic evaluations.

The step-wise approach was employed for CroptimizR to support calibration through key processes of crop growth simulations (Wallach et al., 2025). The value of AIC was determined to incorporate different types of observation data

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

including heading dates, biomass, and yield in order to determine the overall quality of the given parameter set (Table 2). Candidate parameters were assigned to examine the impact of a specific parameter on the overall accuracy. This step is repeated for phenology, biomass, grain number, and yield formation. Simulations of crop growth and development under a wide range of conditions ensure that the calibration process could account for variability in field conditions and avoid overfitting to a single dataset.

When CroptimizR was used for the parameter calibration, the crop growth simulations were repeated until either the change in every parameter became smaller than a specific value unless the number of evaluations exceeded a specific limit. In the present study, the default values of CroptimizR were used to set these stopping criteria. For example, calibration stopped when parameter changes fell below a relative tolerance of 10^{-4} or after 50,000 iterations.

Table 2. List of parameters used for calibration

Group	Major parameter	Candidate parameter
Phenology	P2R, PHINT, P5	P1, P2O, TCLDP
Biomass	G3	-
Grain number	G1	-
Yield	G2	THOT, TCLDF

2.4. Impact of initial values on the local optimization

In conventional calibration protocols, multiple starting points are typically introduced to mitigate convergence to local minima (Wallach et al., 2024). For example, local optimization is repeated using 5 or 20 alternative starting points in addition to the initial parameter set specified by the user for CroptimizR. While this strategy improves the robustness of the optimization procedure, it is not designed to quantify uncertainty arising from the choice of initial values. Moreover, such a limited number of starting points may provide insufficient coverage as the parameter space increases. Therefore, in this study, a large number of random initial parameter sets were generated and used to quantify the impact of the initial values on the reliability of calibrated parameters.

A set of initial parameters $\boldsymbol{P_i}$ were generated to assess the sensitivity of the local optimization method to the starting values using the CroptimizR package. These values were derived from the values of $\boldsymbol{P_o}$ to examine the impact of the distance between the true parameter values and the initial values on the reliability of calibration outcomes. Such intentional variation

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

in initial values facilitated comprehensive exploration of the search space. This also allowed for evaluating whether the proximity of initial guesses to original parameter values could improve the reliability of calibration.

The Manhattan distance was determined between parameter sets due to its effectiveness in multidimensional spaces (Thudumu et al., 2020) and its ability to maintain relative distances between near and far points as dimensionality increases, compared to Euclidean distance (Aggarwal et al., 2001). The Manhattan distance $D_{a,b}$ between two parameter sets, e.g., $\boldsymbol{a}$ and $\boldsymbol{b}$, was determined as follows:

$$D_{a,b} = \sum_{i=1}^{n} |a_i - b_i| \quad (1),$$

where n represents the number of parameters, and a_i and b_i denote the values of the i-th parameter in $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively.

Given the heterogeneous scaling across different parameters, the values of $\boldsymbol{a}$ and $\boldsymbol{b}$ were normalized prior to distance calculation. The search space defined within the parameter input file included in DSSAT was used to set the minimum and the maximum values for each parameter (see Supplementary Information 2). All distance computations were performed using the *philentropy* R package (Drost, 2018), which offers implementations of various distance metrics for multidimensional data analysis.

In the present study, 10,000 sets of $\boldsymbol{P}_i$ values were sampled for calibration. Five bins were defined based on the Manhattan distance between $\boldsymbol{P}_o$ and $\boldsymbol{P}_i$. For each $\boldsymbol{P}_i$ generated, corresponding D values were calculated to assign it to the appropriate bin. When a bin reached its predefined capacity, subsequent $\boldsymbol{P}_i$ values assigned to that bin were discarded. This stratified sampling procedure continued iteratively until all bins were filled with their required number of samples, ensuring a relatively uniform distribution of parameter sets across the feasible space.

Although the D values could range from 0 to 11, bins were set from 0.5 to 5.5, assuming initial values would be selected outside extreme distances. At first, the size of the bin was defined to have 2,000 samples. A large pool of candidate $\boldsymbol{P}_i$ values, equal to 200 times the target number of samples, was generated using a Sobol sequence to ensure sufficient coverage across the parameter space (Zhao et al., 2014). The Sobol sequence was generated using the R package *randtoolbox* (Dutang et al., 2024).

ANOVA and Tukey's honestly significant difference (HSD) tests were performed to examine whether the distance between $\boldsymbol{P}_o$ and $\boldsymbol{P}_i$ would result in the differences in reliability of parameter estimation. The outcomes of parameter calibration were denoted by $\boldsymbol{P}_e$. The Manhattan distance between $\boldsymbol{P}_o$ and $\boldsymbol{P}_e$ or $D_{o,e}$ was used to indicate reliability of the estimates. The parameter estimates may suffer from collinearity between parameters. Following the computation of the $D_{o,e}$ values, the correlation between each pair of parameters was analyzed. The distribution of parameter estimates was also compared by each pair of parameters.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

2.5. Local optimization and global oriented framework

The local optimization and global oriented (LOGO) framework was designed under the assumption that calibration data would be available from multiple sites (Table 3). The reliability of parameter estimates would depend on the number of experimental sites where the observation data were available. Using data from multiple and geographically diverse sites can enhance the spatial portability of the model and reduce the risk of site-specific overfitting (Peng et al., 2020). Multi-site calibration allows for evaluation across a broader range of climatic, soil, and management conditions, enabling more robust estimation of cultivar-specific traits (Lamsal et al., 2017; Kang et al., 2023; Odusanya et al., 2019). Furthermore, it was assumed that crop yield and phenology data were available to meet the minimum data quality standards suggested by Boote et al. (2015).

Table 3. Pseudocode of the Local Optimization Global Oriented (LOGO) calibration framework. The algorithm was implemented in R, and its source code is publicly available on GitHub (https://github.com/EcoInfoLab/LOGO).

Algorithm : LOGO calibration framework

```
Generate calibration datasets from site and initialization parameter set combinations.
for each initial parameter set i:
    for each site combination j:
        Prepare calibration workspace.
        Calibrate cultivar parameters using local optimization.
        Save calibrated parameter set P_ij.
    end
end
Generate a cultivar file using calibrated parameter sets.
for each calibrated parameter set P_ij:
    Run simulations for independent validation sites.
    Save simulated outputs.
end
Compare simulated and observed values.
Estimate posterior parameter distributions and final parameter set.
```

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

At first, the local optimization is performed using the observation data at combinations of site subsets instead of all the sites. In the present study, calibration was performed using data from two to five sites, which would provide a representative range of spatial variability. In this context, calibration using two sites represented a scenario of minimal data availability whereas calibration using five sites approximated an upper bound for the calibration procedures in the present study. The outcome of local optimization using the subset of pseudo-observation data was compared with the parameter estimates obtained from using all the sites. The distance between true and estimated parameter sets was also compared by the number of experiment sites for the local optimization.

The initial guesses of the parameters can be obtained from prior knowledge or values reported in previous studies. It has been recommended to draw initial guesses for new cultivars from the parameter values of cultivars that have similar ecological characteristics or the best parameters from the previous study (Wallach et al., 2024). For example, Yuan et al. (2022) used parameter values from the rice cultivar IR72 as starting values for calibrating other cultivars. In the present study, the $\boldsymbol{P_i}$ values were set using those reported for the rice cultivar IR64 during the second phase of analysis. In addition to the parameters of IR64, five sets of initial values were selected by identifying values nearest to the median distance within each bin of the $D_{o,i}$ values. These initial parameter values were chosen under the assumption that no priori was available due to development of new models or novel parameters for which no previous reference cultivar exists. In total, six calibration sets were obtained for each combination of calibration sites.

The global calibration phase involved the Generalized Likelihood Uncertainty Estimation (GLUE) method to derive the final parameter estimates. While the conventional GLUE approach relies on randomly generated parameter sets to construct the prior distribution, the LOGO algorithm adopted a more targeted strategy. Specifically, the parameter sets were obtained through local optimization across various initial values and site combinations instead of random sampling. This approach allowed for informed construction of prior distribution for the parameter space by leveraging results from the local optimization.

The joint likelihood of each prior parameter set was determined using an independent dataset, which consists of the data from remaining sites not involved in the initial calibration. Likelihood values for crop yield were first computed for each site-year as follows:

$$L_i = \frac{1}{\sqrt{2 \cdot pi \cdot Var}} \cdot exp\left(-\frac{(Obs - Sim)^2}{2 \cdot Var}\right) \quad (2),$$

where *Obs* and *Sim* represent the observed and simulated yield of each site-year, and *Var* represents the variance of the observed yields. The joint likelihood for the parameter set was calculated by multiplying these likelihood values across all the available data, reflecting its overall performance across diverse environments. The probability value (p_i) for each parameter set was calculated as follows:

$$p_i = \frac{L_i}{\sum_{j=1}^{N} L_j} \quad (3),$$

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

where N is the number of the candidate parameters and L_i is the joint likelihood of the i-th parameter set. Finally, the pairs of the parameter set (i) and the probability value (p_i) were used to determine the posterior mean ($\bar{\theta}$) as follows:

$$\bar{\theta} = \sum_{i=1}^{N} p_i \cdot \theta_i \quad (4).$$

The final parameter set represents weighted averages of the most reliable parameter sets across multiple site-years. This approach leverages spatial variability in the data to improve spatial portability and reduce uncertainty in cultivar parameter estimation.

2.6. Evaluation of parameter estimates

The reliability of parameter estimates obtained from the LOGO algorithm was assessed by comparing model outputs with the pseudo-observation data. In particular, heading dates and crop yield simulated using $\boldsymbol{P_e}$ were compared with those of pseudo-observation data from the period of 1991–2000. For evaluation, an independent dataset for the period from 2001 to 2010 was also used to calculate error statistics. This two-period approach enabled assessment of both the consistency of parameter estimation and the predictive accuracy of the calibrated model.

Mean absolute error (MAE) and normalized root mean square error (NRMSE) was calculated to represent the error statistics for heading date and crop yield, respectively. The simulated outputs of heading date and yield were compared by site-year and summarized for the given periods. The values of MAE and NRMSE were determined as follows:

$$\mathrm{MAE} = \frac{\sum |PO - SO|}{n} \quad (5),$$

$$\mathrm{NRMSE} = \sqrt{\frac{\sum (PO - SO)^2}{n}} / \frac{\sum PO}{n} \cdot 100 \quad (6),$$

where n represents the total number of pseudo-observations (PO) and simulation outputs (SO). R^2 was also determined to assess the proportion of variance in the observed data that could be explained by the model simulation.

3. Results

3.1. Impact of initial guesses on local optimization

The parameter estimates were likely similar to the true parameter values when the initial guesses were near the true values (Fig. 2). As the $\boldsymbol{P_i}$ values deviated further from the $\boldsymbol{P_o}$ values, the distance between $\boldsymbol{P_o}$ and $\boldsymbol{P_e}$, $D_{o,e}$ increased accordingly. In particular, an ANOVA followed by Tukey's HSD test indicated that the calibration outcomes were significantly different between the bins based on the magnitude of the distance between $\boldsymbol{P_o}$ and $\boldsymbol{P_i}$, $D_{o,i}$ ($p < 0.05$).

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

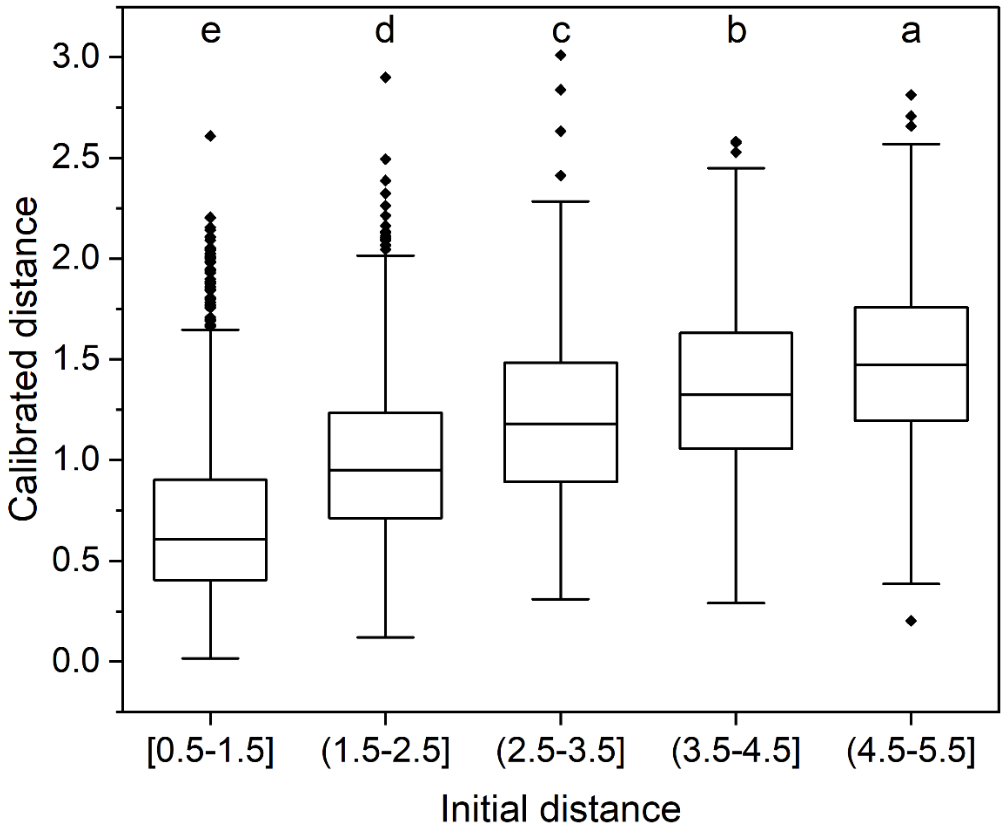

Figure 2: The Manhattan distance between true and estimated parameter values by the bin of the Manhattan distance between true and initial guesses.

The reliability of calibration outcomes differed by the category of cultivar parameters (Fig. 3). Phenology-related parameters, including P2R, PHINT, P5 and P1, had the smallest distance between true and estimated parameter values. In particular, the $D_{o,e}$ values for these parameters remained relatively small across a wide range of the $\boldsymbol{P_i}$ values. For example, the parameter values for PHINT and P5 were estimated within ±5 % of their true values with a probability of 84 % and 92 %, respectively. A similar level of robustness was observed for yield component parameters such as G1 and G2, which were also estimated within ±5 % of their true values at relatively high probabilities, e.g., 72 % and 62 %, respectively. In contrast, the parameter related to grains, e.g., G3, had lower reliability, with the probability of estimation within ±5 % of the true value was 57 %. Stress response parameters such as TCLDP and TCLDF had the greatest variability in $D_{o,e}$, which resulted in generally low reliability as their estimates either remained far from the $\boldsymbol{P_o}$ values or showed minimal deviation from the $\boldsymbol{P_i}$ values.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

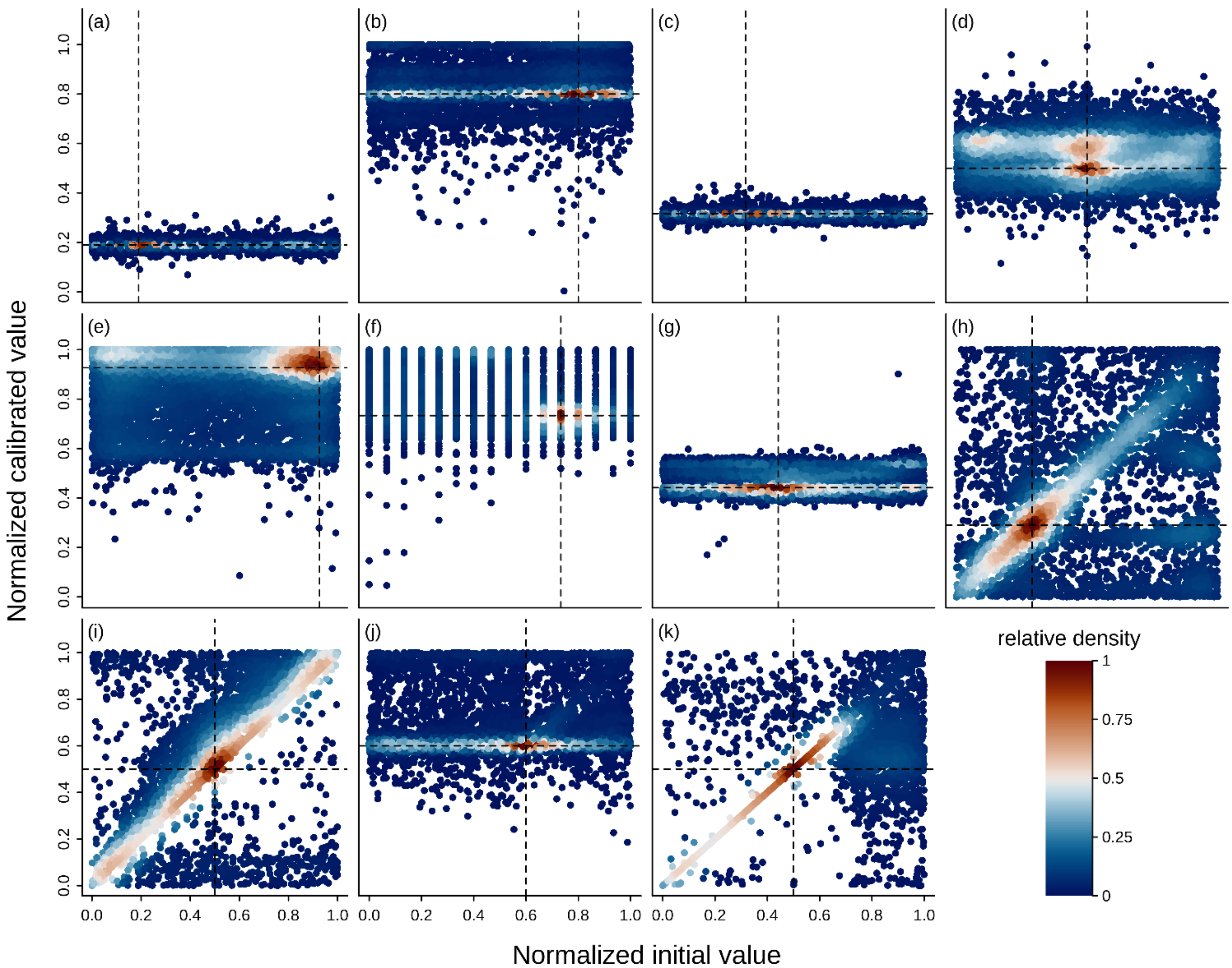

Figure 3: Comparison between initial and calibrated parameters for (a) P2R, (b) PHINT, (c) P5, (d) G3, (e) G1, (f) G2, (g) P1, (h) P2O, (i) TCLDP, (j) THOT, and (k) TCLDF. Dotted lines represent the true parameter.

Most pairs between parameter estimates had significant correlations when a wide range of initial guesses were used (Fig. 4). Although many of the correlation coefficients r were relatively small, several exceptions indicated strong associations among calibrated parameters. Such relationships were particularly pronounced within the same parameter categories. For example, P1 and P2O had a strong positive correlation (r = 0.86) while G1 and G2 had a strong negative correlation (r = -0.68). Meaningful correlations were also observed across categories, e.g., P1 and G3 (r = 0.57), and PHINT and G3 (r = 0.58). In particular, P2O had a relatively large number of pronounced pairwise correlations, second only to G3, which are less likely to be determined from field observations. Furthermore, the average coefficient of determination for the parameter pairs involving P2O was higher (0.159) than that for the other parameters.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

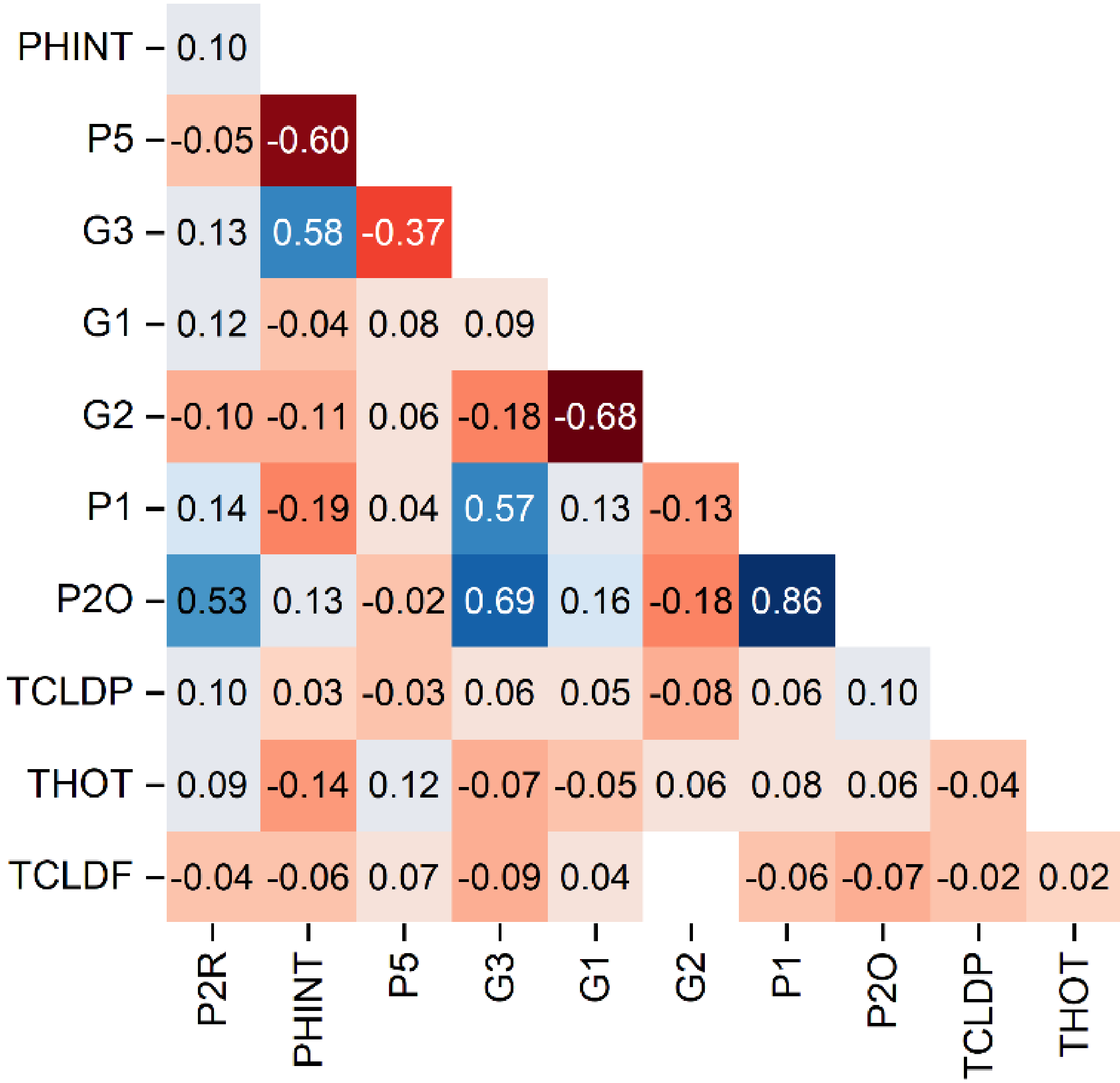

Figure 4: Correlation matrix among the calibrated parameters. Only significant Pearson correlation coefficients ($p < 0.05$) are shown.

Errors in crop growth simulations tended to increase when cultivar parameters were estimated from initial guesses that had greater distance from the true parameter values (Fig. 5). Differences in heading date and yield varied significantly by the distance between $\boldsymbol{P_o}$ and $\boldsymbol{P_i}$ ($p < 0.05$). In addition, the median error values increased with larger $D_{o,i}$ values. Nevertheless, the overall magnitude of the errors remained relatively small, e.g., < 5 %, even when the distance between true and estimated parameters was relatively large, e.g., the $D_{o,e}$ value > 4.5 (see supplementary information 3).

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

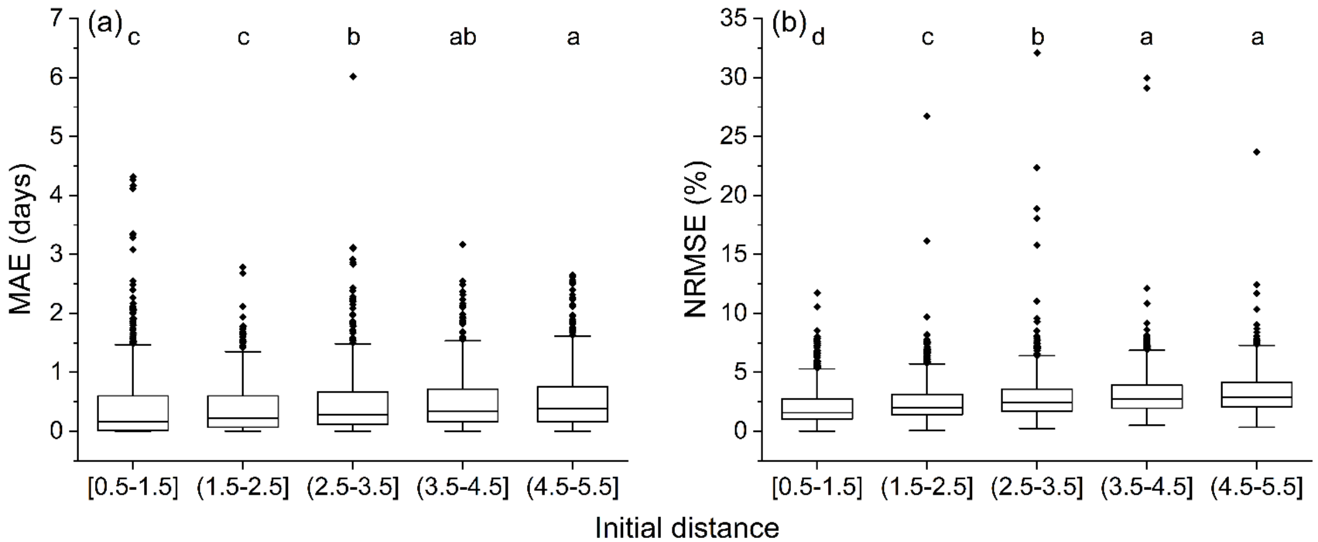

Figure 5: Distribution of prediction errors for (a) heading date and (b) yield by the group of the Manhattan distance between true parameter values and their initial guess.

3.2. Outcome of local optimization by combination of sites

There was a large variation in the reliability of parameter estimates by the number of sites and initial values used for the local optimization method (Fig. 6). Overall, the value of $D_{o,e}$ tended to decrease with the greater number of site combinations where the observation data were obtained during the calibration. However, such a tendency was inconsistent by the initial guess of the parameter values. For example, the estimates of parameters were deviated from the true parameters even when the value of $D_{o,i}$ was less for a set of initial values than that for IR64.

The use of data gathered at all the experiment sites did not necessarily result in more reliable parameter sets using the local optimization method (Fig. 6). The distance between true and estimated parameters differed by the initial values of the calibration when pseudo-observation data from all the sites were used. In contrast, there were specific combinations of sites where the $D_{o,e}$ values were relatively smaller than the other sites. For example, combinations of three sites had the smallest value of $D_{o,e}$ when the parameter values for the IR64 was used for the initial guesses. Still, the reliability of parameter estimates derived from the combinations of sites differed by the initial values.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

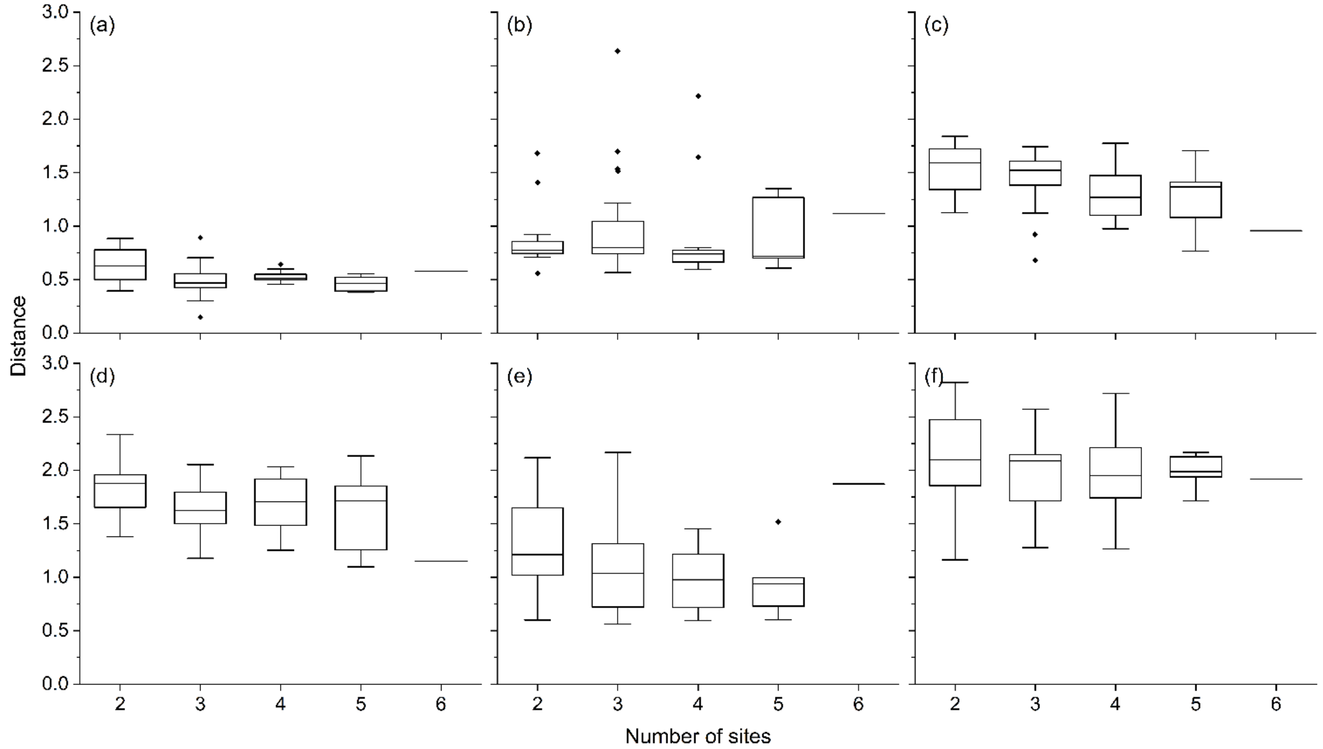

Figure 6: Distance between the true parameter and calibrated parameter sets under the different numbers of site combinations. Panels a-f represent results for calibration using (a) IR64 and (b-f) random parameter sets. The distance between the true parameter and initial parameter was (a) 1.68, (b) 1.04, (c) 2.01, (d) 3.01, (e) 3.99, and (f) 5.34, respectively.

3.3. Outcome of LOGO algorithm

The implementation of the LOGO procedure resulted in relatively smaller distances between the true and estimated parameter values compared with those obtained from local optimization using all the sites under the assumption that no IR64 was available for the initial guess (Fig. 7). This improvement was consistent across most cases, regardless of the number of sites used or the distance from the initial parameter set. For example, the $D_{o,e}$ value derived from the LOGO method was located near the first quantile of the distribution of $D_{o,e}$ values obtained from local optimization with the site combinations. In particular, parameter estimates using more than four site combinations had lower $D_{o,e}$ values using the LOGO algorithm.

The parameter set derived from the LOGO algorithm had strong agreement between simulated data and pseudo-observation during the evaluation (Fig. 8). For example, simulated heading dates and yield closely matched the observations using the parameter set obtained from four site combinations. The regression slopes were close to 1 and the R^2 values exceeded 0.99 for both variables because observation data were derived from the model itself. Using different numbers of sites for calibration also resulted in similar accuracy for all variables in the evaluation sets (see supplementary information 4).

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

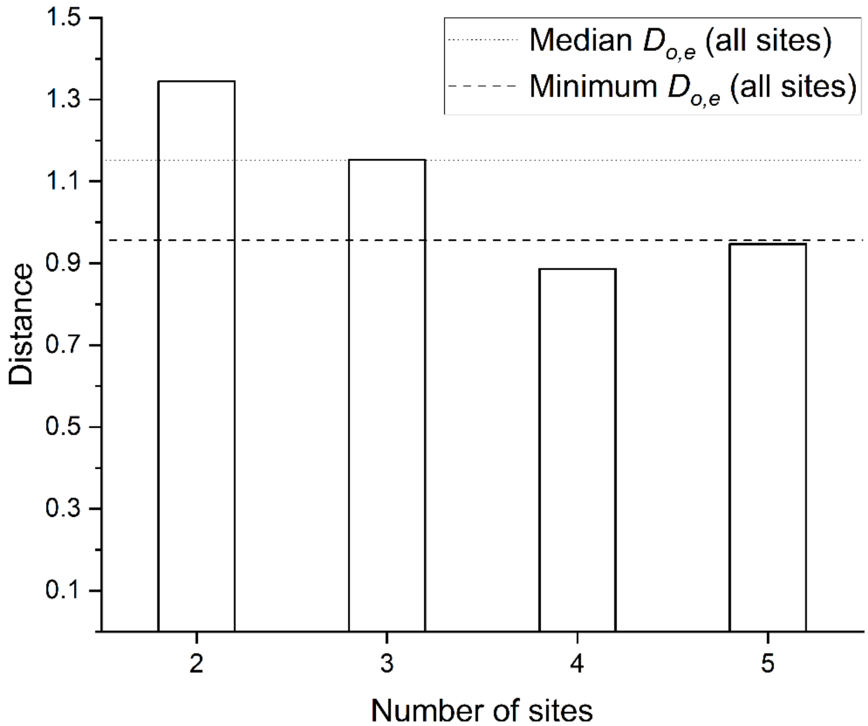

Figure 7: Effect of the GLUE procedure on the distance between the calibrated and true parameter sets, using all data from the calibration with five different random initial values. Dashed and dotted lines represent the minimum and median values of the distances when all sites are used for calibration.

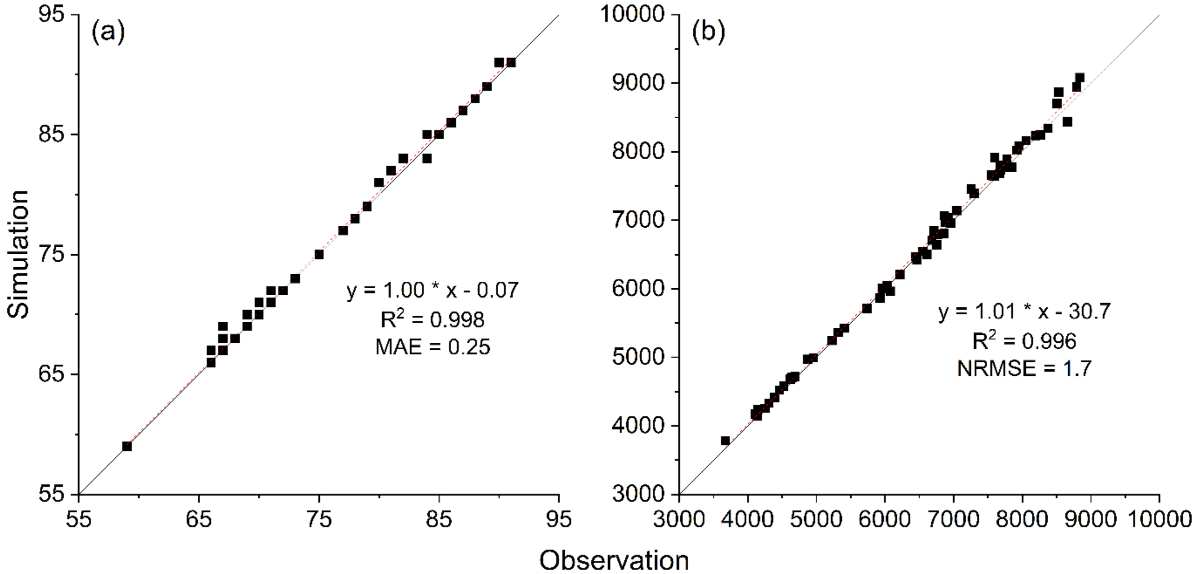

Figure 8: Evaluation results for (a) heading dates (days after planting) and (b) yield (kg ha^{-1}) using the parameter set derived from the LOGO algorithm using four site combinations.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

4. Discussion

Our results demonstrated that the LOGO approach allowed for calibration of parameters close to true parameters taking advantage of local search and uncertainty assessment for distributed experimental sites. In particular, the use of several initial values supported robust estimation of parameters. Although there was no guarantee such that the LOGO algorithm would result in the maximum reliability, it provided parameter sets with relatively smaller distance between true and estimated parameters without prior knowledge of the true values for the parameters. Such a feature can minimize the risk of equifinality, a condition in which distinct parameter sets result in similar crop growth simulations, thereby obscuring the identification of true parameter values (Beven, 2019).

Our calibration outcomes highlighted that parameters are not necessarily equally identifiable using a single set of initial guesses. Key phenology parameters were estimated with relatively high consistency over a wide range of initial values in the present study. These parameters had smaller deviations from their known true values, indicating that the model outputs were strongly influenced by them and thus provided clear signals for calibration. This aligns with prior findings that phenology-related genotype coefficients were often constrained by observed flowering and maturity dates (Pathak et al., 2012). In contrast, yield-component parameters, e.g., those governing potential grain number or individual grain weight, had lower identifiability as their estimated values varied widely and often deviated further from the true values. The larger variability of these parameters suggests that their multiple combinations could produce similar yield predictions due to compensatory effects in the model. This is a classic symptom of equifinality, where different parameter sets yield equivalent model outputs (Dzotsi et al., 2015). In addition, some parameters may be insensitive under the given conditions, which makes it challenging to calibrate them reliably (Chapagain et al., 2022).

Danilova et al. (2020) suggested that different initializations can yield different calibrated parameter sets, underscoring the non-convex nature of model calibration problems. We found that the choice of initial parameter values could significantly influence calibration outcomes. When the initial guess for a parameter was very close to the true value, one might expect that the calibration would converge to the correct value. Indeed, it was observed that parameters initialized near their true value tended to have accurate parameter estimates. Nevertheless, proximity alone was not a guarantee of successful parameter calibration. In fact, it was also found that the final distance between true and estimated value was larger than the initial distance although such events occurred intermittently. This counter-intuitive outcome where the calibration moved the parameter estimates further from the truth indicated the presence of local minima or complex error surfaces. An initial guess that is distant from true values can also lead the algorithm to a spurious local optimum if the search becomes trapped (Tan et al., 2019).

The parameter set that had moderate deviations from the true values, yet remained applicable to an actual cultivar, led to reasonable calibration outcomes in the present study. These findings align with the broader understanding that local optimization methods commonly used in crop modelling, e.g., Levenberg–Marquardt algorithms (Seidel et al., 2018), often start from the parameter values of a known cultivar. When such a cultivar was not available, global optimization and

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

sampling-based methods such as Genetic Algorithms and GLUE have been applied to reduce dependence on starting conditions (Li et al., 2018; Tan et al., 2019). The evidence of initial-value effects underscores the importance of initiating calibration from multiple starting points or alternatively integrating local optimization with global search strategies (see Supplementary Information 5), to minimize the risk of obtaining parameter estimates that are artifacts of suboptimal initial guesses.

Analysis of the calibrated parameter sets indicated interactions and trade-offs among parameters. It was found that there was a clear correlation between the calibrated values of the grain filling parameter (G3) and the phyllochron interval (PHINT). When one of these parameters became high in a given calibration, the other tended to be low, suggesting a compensatory relationship. In the crop model, a faster leaf appearance rate, e.g., shorter PHINT, could be offset by a slower grain filling rate, e.g., lower G3, to produce a similar duration from planting to maturity and final yield, and vice versa. The model appears to exploit this trade-off to fit the data, highlighting a case of parameter interaction where two parameters can adjust in tandem to buffer the output. This kind of inverse correlation indicates the equifinality in complex models such that multiple parameter combinations yield equivalent predictions (Dzotsi et al., 2015). The use of the LOGO algorithm made it possible to avoid such problems by applying a global method followed by the local optimization from the multiple starting values.

Kawakita et al. (2024) proposed a dual ensemble framework that combines multiple rice phenology models with machine-learning regression models trained on genetic information to predict cultivar-specific parameters and heading dates. In their study, the calibrated parameters were first obtained from conventional fitting and then treated as reference values for training the genotype–parameter regression. This approach does not eliminate the equifinality issue inherent to calibration since the reference parameters still originate from model fitting. However, the use of genotypic prediction rather than phenotypic fitting can mitigate the dependence on initial values in crop model calibration as genetic markers would provide informative priors for parameter estimation.

One of the limitations in the present study was the use of pseudo-observation data, although the proposed method was evaluated through a simple case study (see Supplementary Information 6). It was possible to demonstrate the impact of initial values on the outcome of calibration using the pseudo-observation data because no uncertainty was embedded in the observation data. Still, uncertainty and availability of data can affect the reliability of calibration. The field observations for calibration data are often exposed to measurement errors, e.g., yield measurement inaccuracies and weather data errors, which may not represent the true field conditions due to unobserved factors. Such noise in the data can mislead the outcome of calibration. For example, an outlier yield observation could pull the parameter estimation in a direction that degrades performance elsewhere.

The influence of uncertainty inherent in the observation data on the LOGO framework can be further examined by introducing controlled levels of random error into pseudo-observation datasets in future studies. By gradually increasing the magnitude and structure of noise, it would be possible to characterize the error tolerance of the calibration process and to determine the minimum data quality required for reliable parameter estimation. Following this controlled assessment,

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

applying the LOGO procedure to real field datasets across diverse environments could verify its robustness under practical variations in management and observation quality. In addition, extending the framework to other cultivars and crop species would help determine whether the hybrid optimization strategy consistently mitigates equifinality and initial-value sensitivity across different physiological systems (see Supplementary Information 5). Such efforts will enhance the general applicability of the LOGO approach and support the development of more transferable and reproducible calibration protocols.

5. Conclusion

Calibration outcomes in process-based crop growth models can be strongly affected by the choice of initial parameter values, leading to equifinality and convergence to local minima. This study demonstrates that such sensitivity was mitigated by integrating local and global optimization procedures within the LOGO framework. This improvement reflects the systematic exploration of multiple initial values and site combinations, which reduced equifinality and improved parameter robustness across environments. Still, because this study utilized pseudo-observation datasets to isolate the role of calibration procedures, future research should explicitly account for data uncertainty and measurement errors that typify real field observations.

Code and data availability

Crop growth simulations were conducted using the CSM-CERES-Rice model included in the Decision Support System for Agrotechnology Transfer (DSSAT, version 4.8.2). DSSAT was compiled from publicly available source code (https://github.com/DSSAT/dssat-csm-os).

Parameter calibration was performed using the R package croptimizR, which is publicly available.

Climate data were derived from the AgERA5 dataset, which is freely available under the CC BY license, and were obtained from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/datasets/sis-agrometeorological-indicators).

Soil data were derived from the SoilGrids database, which is publicly available under the CC BY 4.0 license (https://soilgrids.org/).

Climate and soil data used in Supplementary Information 6 were obtained from weather stations operated by the Korean Meteorological Administration (https://data.kma.go.kr) and soil series data provided by the Rural Development Administration (https://soil.rda.go.kr).

The current version of the LOGO framework and sample input files is available from the project website https://github.com/EcoInfoLab/LOGO under the MIT licence. The exact version of the LOGO implementation used to produce the results presented in this paper is archived on Zenodo under https://doi.org/10.5281/zenodo.20707899 (Hyun et al., 2026a).

All model input and output datasets used in this study are archived on Figshare (https://doi.org/10.6084/m9.figshare.32209782), including management, pseudo-observations, initial parameter sets, calibrated parameter sets, and simulation outputs (Hyun et al., 2026b).

Author contributions

Shinwoo Hyun (Methodology, Data curation, Writing - original draft), Junhwan Kim (Conceptualization, Resources, Writing - original draft), Young Sang Joh (Data curation), Kwang Soo Kim (Conceptualization, Methodology, Writing - original draft), Gerrit Hoogenboom (Writing - review & editing)

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

During the preparation of this manuscript, the authors used ChatGPT for language editing to improve clarity and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the final content of the published article.

Financial support

This work was supported by Cooperative Research Program for Agriculture Science & Technology Development (Project No. RS-2025-02232989), funded by Rural Development Administration, Republic of Korea. This work was also supported by “Statistics from Space: Next-Generation Agricultural Production Information for Enhanced Monitoring of Food Security in Mozambique”, funded by Ministry of Agriculture, Food and Rural Affairs.

References

Aggarwal, C. C., Hinneburg, A., and Keim, D. A.: On the surprising behavior of distance metrics in high dimensional space, International conference on database theory, 420-434, https://doi.org/10.1007/3-540-44503-x_27, 2001.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

Archontoulis, S. V., Miguez, F. E., and Moore, K. J.: A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM PLANT model: application to soybean, Environmental modelling & software, 62, 465-477, https://doi.org/10.1016/j.envsoft.2014.04.009, 2014.

Archontoulis, S. V., Castellano, M. J., Licht, M. A., Nichols, V., Baum, M., Huber, I., Martinez-Feria, R., Puntel, L., Ordóñez, R. A., and Iqbal, J.: Predicting crop yields and soil-plant nitrogen dynamics in the US Corn Belt, Crop Science, 60, 721-738, https://doi.org/10.1002/csc2.20039, 2020.

Beck, H. E., McVicar, T. R., Vergopolan, N., Berg, A., Lutsko, N. J., Dufour, A., Zeng, Z., Jiang, X., Van Dijk, A. I., and Miralles, D. G.: High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections, Scientific data, 10, 724, https://doi.org/10.1038/s41597-023-02549-6, 2023.

Beven, K.: Validation and equifinality, in: Computer simulation validation: Fundamental concepts, methodological frameworks, and philosophical perspectives, Springer, 791-809, https://doi.org/10.1007/978-3-319-70766-2_32, 2019.

Boogaard, H., van der Grijn, G., and de Wit, A.: Agrometeorological indicators from 1979 to present derived from reanalysis, 2020.

Boote, K. J., Porter, C., Jones, J. W., Thorburn, P. J., Kersebaum, K., Hoogenboom, G., White, J., and Hatfield, J.: Sentinel site data for crop model improvement—definition and characterization, Improving modeling tools to assess climate change effects on crop response, 7, 125-158, https://doi.org/10.2134/advagricsystmodel7.2014.0019, 2016.

Buis, S., Lecharpentier, P., Vezy, R., and Giner, M.: CroptimizR: A Package to Estimate Parameters of Crop Models (v0.6.1), Zenodo [code], https://doi.org/10.5281/zenodo.10425387, 2023.

Chapagain, R., Remenyi, T. A., Harris, R. M., Mohammed, C. L., Huth, N., Wallach, D., Rezaei, E. E., and Ojeda, J. J.: Decomposing crop model uncertainty: A systematic review, Field Crops Research, 279, 108448, https://doi.org/10.1016/j.fcr.2022.108448, 2022.

Danilova, M., Dvurechensky, P., Gasnikov, A., Gorbunov, E., Guminov, S., Kamzolov, D., and Shibaev, I.: Recent theoretical advances in non-convex optimization, in: High-Dimensional Optimization and Probability: With a View Towards Data Science, Springer, 79-163, https://doi.org/10.1007/978-3-031-00832-0_3, 2022.

Dou, Z., Li, Y., Guo, H., Chen, L., Jiang, J., Zhou, Y., Xu, Q., Xing, Z., Gao, H., and Zhang, H.: Effects of mechanically transplanting methods and planting densities on yield and quality of Nanjing 2728 under rice-crayfish continuous production system, Agronomy, 11, 488, https://doi.org/10.3390/agronomy11030488, 2021.

Drost, H.-G.: Philentropy: information theory and distance quantification with R, Journal of Open Source Software, 3, 765, https://doi.org/10.21105/joss.00765, 2018.

Dutang, C., Savicky, P., Chalabi, Y., Wuertz, D., Knuth, D., Matsumoto, M., Saito, M., and Dutang, M. C.: Package 'randtoolbox', https://doi.org/10.32614/cran.package.randtoolbox, 2024.

Dzotsi, K., Basso, B., and Jones, J.: Parameter and uncertainty estimation for maize, peanut and cotton using the SALUS crop model, Agricultural Systems, 135, 31-47, https://doi.org/10.1016/j.agsy.2014.12.003, 2015.

https://doi.org/10.5194/egusphere-2026-3538
Preprint. Discussion started: 28 July 2026

Ferreira, T. B., Shelia, V., Porter, C., Cadena, P. M., Cortasa, M. S., Khan, M. S., Pavan, W., and Hoogenboom, G.: Enhancing crop model parameter estimation across computing environments: Utilizing the GLUE method and parallel computing for determining genetic coefficients, Computers and Electronics in Agriculture, 227, 109513, https://doi.org/10.1016/j.compag.2024.109513, 2024.

Gao, Y., Wallach, D., Hasegawa, T., Tang, L., Zhang, R., Asseng, S., Kahveci, T., Liu, L., He, J., and Hoogenboom, G.: Evaluation of crop model prediction and uncertainty using Bayesian parameter estimation and Bayesian model averaging, Agricultural and Forest Meteorology, 311, 108686, https://doi.org/10.1016/j.agrformet.2021.108686, 2021.

Gao, Y., Wallach, D., Liu, B., Dingkuhn, M., Boote, K. J., Singh, U., Asseng, S., Kahveci, T., He, J., and Zhang, R.: Comparison of three calibration methods for modeling rice phenology, Agricultural and Forest Meteorology, 280, 107785, https://doi.org/10.1016/j.agrformet.2019.107785, 2020.

Gavasso-Rita, Y. L., Papalexiou, S. M., Li, Y., Elshorbagy, A., Li, Z., and Schuster-Wallace, C.: Crop models and their use in assessing crop production and food security: A review, Food and Energy Security, 13, e503, https://doi.org/10.1002/fes3.503, 2024.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., and Bauer-Marschallinger, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLoS one, 12, e0169748, 2017.

Hoogenboom, G., Porter, C., Shelia, V., Boote, K., Singh, U., Pavan, W., Oliveira, F., Moreno-Cadena, L., Ferreira, T., and White, J.: Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.8.2 (www.DSSAT.net). DSSAT Foundation, Gainesville, Florida, USA. https://dssat. net, 2023.

Hoogenboom, G., Porter, C. H., Boote, K. J., Shelia, V., Wilkens, P. W., Singh, U., White, J. W., Asseng, S., Lizaso, J. I., and Moreno, L. P.: The DSSAT crop modeling ecosystem, in: Advances in crop modelling for a sustainable agriculture, Burleigh Dodds Science Publishing, 173-216, 2019.

Hyun, S., Kim, J., Joh, Y. S., Kim, K. S., and Hoogenboom, G.: EcoInfoLab/LOGO: LOGO v1.0.0 (v1.0.0), Zenodo [code], https://doi.org/10.5281/zenodo.20707899, 2026a.

Hyun, S., Kim, J., Joh, Y. S., Kim, K. S., and Hoogenboom, G.: Datasets for calibration and evaluation of the LOGO framework, https://doi.org/10.6084/m9.figshare.32209782.v2, 2026b.

Hyun, S., Park, J. Y., Kim, J., Fleisher, D. H., and Kim, K. S.: GLUEOS: A high performance computing system based on the orchestration of containers for the GLUE parameter calibration of a crop growth model, Computers and Electronics in Agriculture, 197, 106906, https://doi.org/10.1016/j.compag.2022.106906, 2022.

Jacob, P., Robert, C. P., and Smith, M. H.: Using parallel computation to improve independent Metropolis–Hastings based estimation, Journal of Computational and Graphical Statistics, 20, 616-635, https://doi.org/10.1198/jcgs.2011.10167, 2011.

Jeraj, R., Wu, C., and Mackie, T. R.: Optimizer convergence and local minima errors and their clinical importance, Physics in Medicine & Biology, 48, 2809-2827, https://doi.org/10.1088/0031-9155/48/17/306 2003.

Kang, D. G., Kim, Y.-U., Hyun, S., Kim, K. S., Kim, J., Lee, C.-K., Maruyama, A., Beresford, R. M., and Fleisher, D. H.: Identification of a spatial distribution threshold for the development of a solar radiation model using deep neural networks, Environmental Research Letters, 18, 104020, https://doi.org/10.1088/1748-9326/acf6d4 2023.

Kaur, P., Saini, K., Kaur, K., and Kaur, J.: Production potential of short duration rice (Oryza sativa) in relation to age of seedlings, date of transplanting and crop geometry, The Indian Journal of Agricultural Sciences, 93, https://doi.org/10.56093/ijas.v93i8.138424 2023.

Kawakita, S., Yamasaki, M., Teratani, R., Yabe, S., Kajiya-Kanegae, H., Yoshida, H., Fushimi, E., and Nakagawa, H.: Dual ensemble approach to predict rice heading date by integrating multiple rice phenology models and machine learning-based genetic parameter regression models, Agricultural and Forest Meteorology, 344, 109821, https://doi.org/10.1016/j.agrformet.2023.109821 2024.

Kim, K. S., Yoo, B. H., Shelia, V., Porter, C. H., and Hoogenboom, G.: START: A data preparation tool for crop simulation models using web-based soil databases, Computers and Electronics in Agriculture, 154, 256-264, https://doi.org/10.1016/j.compag.2018.08.023 2018.

Kim, J., Sang, W., Shin, P., Baek, J., Kwon, D., Lee, Y., Cho, J.-I., and Seo, M.: Long-term monitoring data for growth and yield of local rice varieties in South Korea, Korean Journal of Agricultural and Forest Meteorology, 22, 176-182, https://doi.org/10.5532/kjafm.2020.22.3.176, 2020.

Lamsal, A., Welch, S., Jones, J., Boote, K., Asebedo, A., Crain, J., Wang, X., Boyer, W., Giri, A., and Frink, E.: Efficient crop model parameter estimation and site characterization using large breeding trial data sets, Agricultural systems, 157, 170-184, https://doi.org/10.1016/j.agsy.2017.07.016 2017.

Li, Z., He, J., Xu, X., Jin, X., Huang, W., Clark, B., Yang, G., and Li, Z.: Estimating genetic parameters of DSSAT-CERES model with the GLUE method for winter wheat (Triticum aestivum L.) production, Computers and electronics in agriculture, 154, 213-221, https://doi.org/10.1016/j.compag.2018.09.009 2018.

Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Adam, M., Bregaglio, S., Buis, S., Confalonieri, R., and Fumoto, T.: Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions, Global change biology, 21, 1328-1341, https://doi.org/10.1111/gcb.12758 2015.

Nelder, J. A. and Mead, R.: A simplex method for function minimization, The computer journal, 7, 308-313, https://doi.org/10.1093/comjnl/7.4.308 1965.

Odusanya, A. E., Mehdi, B., Schürz, C., Oke, A. O., Awokola, O. S., Awomeso, J. A., Adejuwon, J. O., and Schulz, K.: Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria, Hydrology and Earth System Sciences, 23, 1113-1144, https://doi.org/10.5194/hess-23-1113-2019 2019.

Pathak, T. B., Jones, J. W., Fraisse, C. W., Wright, D., and Hoogenboom, G.: Uncertainty analysis and parameter estimation for the CSM-CROPGRO-Cotton model, Agronomy Journal, 104, 1363-1373, https://doi.org/10.2134/agronj2011.0349 2012.

Peng, S., Cassman, K. G., Virmani, S. S., Sheehy, J., and Khush, G. S.: Yield Potential Trends of Tropical Rice since the Release of IR8 and the Challenge of Increasing Rice Yield Potential, Crop Science, 39, 1552-1559, https://doi.org/10.2135/cropsci1999.3961552x, 1999.

Peng, S., Laza, R. C., Visperas, R. M., Sanico, A. L., Cassman, K. G., and Khush, G. S.: Grain Yield of Rice Cultivars and Lines Developed in the Philippines since 1966, Crop Science, 40, 307-314, https://doi.org/10.2135/cropsci2000.402307x, 2000.

Peng, B., Guan, K., Tang, J., Ainsworth, E. A., Asseng, S., Bernacchi, C. J., Cooper, M., Delucia, E. H., Elliott, J. W., Ewert, F., Grant, R. F., Gustafson, D. I., Hammer, G. L., Jin, Z., Jones, J. W., Kimm, H., Lawrence, D. M., Li, Y., Lombardozzi, D. L., Marshall-Colon, A., Messina, C. D., Ort, D. R., Schnable, J. C., Vallejos, C. E., Wu, A., Yin, X., and Zhou, W.: Towards a multiscale crop modelling framework for climate change adaptation assessment, Nature Plants, 6, 338-348, https://doi.org/10.1038/s41477-020-0625-3, 2020.

Potgieter, A. B., Zhao, Y., Zarco-Tejada, P. J., Chenu, K., Zhang, Y., Porker, K., Biddulph, B., Dang, Y. P., Neale, T., Roosta, F., and Chapman, S.: Evolution and application of digital technologies to predict crop type and crop phenology in agriculture, in silico Plants, 3, diab017, https://doi.org/10.1093/insilicoplants/diab017, 2021.

Rahman, M. S., Tareq, T. M., Sarker, P. K., Rashid, E. S. M. H., Yasmeen, R., Ali, M. A., Seraj, Z. I., and Shimono, H.: Genetic variation of phenotypic plasticity in Bangladesh rice germplasm, Field Crops Research, 243, 107618, https://doi.org/10.1016/j.fcr.2019.107618, 2019.

Rezaei, E. E., Webber, H., Asseng, S., Boote, K., Durand, J. L., Ewert, F., Martre, P., and MacCarthy, D. S.: Climate change impacts on crop yields, Nature Reviews Earth & Environment, 4, 831-846, https://doi.org/10.1038/s43017-023-00491-0, 2023.

Seidel, S. J., Palosuo, T., Thorburn, P., and Wallach, D.: Towards improved calibration of crop models – Where are we now and where should we go?, European Journal of Agronomy, 94, 25-35, https://doi.org/10.1016/j.eja.2018.01.006, 2018.

Shoarinezhad, V., Wieprecht, S., and Haun, S.: Comparison of Local and Global Optimization Methods for Calibration of a 3D Morphodynamic Model of a Curved Channel, Water, 12, 1333, https://doi.org/10.3390/w12051333, 2020.

Tan, J., Cao, J., Cui, Y., Duan, Q., and Gong, W.: Comparison of the Generalized Likelihood Uncertainty Estimation and Markov Chain Monte Carlo Methods for Uncertainty Analysis of the ORYZA_V3 Model, Agronomy Journal, 111, 555-564, https://doi.org/10.2134/agronj2018.05.0336, 2019.

Thudumu, S., Branch, P., Jin, J., and Singh, J.: A comprehensive survey of anomaly detection techniques for high dimensional big data, Journal of Big Data, 7, 42, https://doi.org/10.1186/s40537-020-00320-x, 2020.

Timlin, D., Paff, K., and Han, E.: The role of crop simulation modeling in assessing potential climate change impacts, Agrosystems, Geosciences & Environment, 7, e20453, https://doi.org/10.1002/agg2.20453, 2024.

Wallach, D., Buis, S., Seserman, D.-M., Palosuo, T., Thorburn, P. J., Mielenz, H., Justes, E., Kersebaum, K.-C., Dumont, B., Launay, M., and Seidel, S. J.: A calibration protocol for soil-crop models, Environmental Modelling & Software, 180, 106147, https://doi.org/10.1016/j.envsoft.2024.106147, 2024.

Wallach, D., Kim, K. S., Hyun, S., Buis, S., Thorburn, P., Mielenz, H., Seidel, S. J., Alderman, P. D., Dumont, B., Fallah, M. H., Hoogenboom, G., Justes, E., Kersebaum, K.-C., Launay, M., Leolini, L., Mehmood, M. Z., Moriondo, M., Jing, Q., Qian, B., Susanne, S., Seserman, D.-M., Shelia, V., Weihermüller, L., and Palosuo, T.: Evaluating the AgMIP calibration protocol for crop models; case study and new diagnostic tests, European Journal of Agronomy, 168, 127659, https://doi.org/10.1016/j.eja.2025.127659, 2025.

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., Jabloun, M., Jansson, P.-E., Jing, Q., Justes, E., Kersebaum, K. C., Klosterhalfen, A., Launay, M., Lewan, E., Luo, Q., Maestrini, B., Mielenz, H., Moriondo, M., Nariman Zadeh, H., Padovan, G., Olesen, J. E., Poyda, A., Priesack, E., Pullens, J. W. M., Qian, B., Schütze, 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., and Seidel, S. J.: The chaos in calibrating crop models: Lessons learned from a multi-model calibration exercise, Environmental Modelling & Software, 145, 105206, https://doi.org/10.1016/j.envsoft.2021.105206, 2021.

Yoo, B. H., Kim, K. S., Kim, J., Beresford, R. M., and Fleisher, D. H.: The development of a soil data aggregation method for the spatial impact assessment of climate change on soybean yield, Computers and Electronics in Agriculture, 235, 110347, https://doi.org/10.1016/j.compag.2025.110347, 2025.

Yuan, S., Stuart, A. M., Laborte, A. G., Rattalino Edreira, J. I., Dobermann, A., Kien, L. V. N., Thúy, L. T., Paothong, K., Traesang, P., Tint, K. M., San, S. S., Villafuerte, M. Q., Quicho, E. D., Pame, A. R. P., Then, R., Flor, R. J., Thon, N., Agus, F., Agustiani, N., Deng, N., Li, T., and Grassini, P.: Southeast Asia must narrow down the yield gap to continue to be a major rice bowl, Nature Food, 3, 217-226, https://doi.org/10.1038/s43016-022-00477-z, 2022.

Zhao, G., Bryan, B. A., and Song, X.: Sensitivity and uncertainty analysis of the APSIM-wheat model: Interactions between cultivar, environmental, and management parameters, Ecological Modelling, 279, 1-11, https://doi.org/10.1016/j.ecolmodel.2014.02.003, 2014.

Zhu, Z.-R., Romena, A. M., and Cohen, M. B.: Comparison of stem borer damage and resistance in semidwarf Indica rice varieties and prototype lines of a new plant type, Field Crops Research, 75, 37-45, https://doi.org/10.1016/S0378-4290(02)00004-7, 2002.