

Article

Calibration and Validation of CSM-CROPGRO-Lentil Under Semi-Arid Conditions: Model Performance and Genotype Ranking Capacity Across Contrasting Drought Seasons

Mustafa Ceritoglu 

Department of Field Crops, Faculty of Agriculture, Siirt University, Siirt 56100, Türkiye; ceritoglu@siirt.edu.tr

Abstract

The CSM-CROPGRO-Lentil model (DSSAT v4.8.5) was calibrated and validated for 43 ICARDA lentil genotypes across three growing seasons (2019–2022) under semi-arid rainfed conditions in southeastern Türkiye. Calibration against the 2019–2020 season yielded excellent performance for phenological traits (anthesis date $nRMSE = 0.8\%$; maturity date $nRMSE = 1.2\%$), thousand-seed weight ($nRMSE = 1.6\%$), and good performance for grain yield ($nRMSE = 19.3\%$; $d = 0.935$) and biological yield ($nRMSE = 11.1\%$; $d = 0.941$). Phenological simulations remained robust (anthesis and maturity dates $nRMSE = 6.8\%$ and 1.6%), but grain and biological yield simulations failed ($nRMSE = 43.0$ – 74.7%) with opposing bias directions in both validation seasons. Spearman rank correlation collapsed to non-significant levels in both validation years ($\rho = -0.020$ and $+0.263$), indicating a loss of genotype-ranking capacity, in which only three accessions (G37, G3695, and G21151) maintained simulation error below 40% across both drought seasons. Model bias was not significantly associated with PEG-derived drought tolerance classifications, based on composite stress tolerance index-based reclassification with balanced group sizes ($n = 5$ vs. $n = 5$; Mann–Whitney $p \geq 0.548$), suggesting that limitations in species-level water stress response functions may have contributed to model failure, although the potential influence of interannual differences in soil hydraulic properties cannot be excluded. Late-maturing genotypes were simulated significantly more accurately than early and mid-maturing genotypes in 2021–2022 (Kruskal–Wallis $p = 0.013$), while large-seeded genotypes showed disproportionately greater underestimation in 2020–2021 ($\rho = -0.410$, $p = 0.006$). These findings demonstrate that CROPGRO-Lentil reliably simulates phenological diversity under favorable conditions but requires genotype-specific drought stress parameterization and improved seed-filling dynamics to support genotype evaluation under increasingly variable semi-arid moisture regimes.



Academic Editors: Lea Piscitelli,
Lorenzo Negri and Mohamed
Houssemeddine Sellami

Received: 18 June 2026

Revised: 22 July 2026

Accepted: 27 July 2026

Published: 29 July 2026

Copyright: © 2026 by the author.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: CSM-CROPGRO-Lentil; DSSAT; genotype ranking; modelling; semi-arid conditions

1. Introduction

Lentil (*Lens culinaris* Medik.) is one of the earliest domesticated legume crops and occupies a central role in global food and nutritional security, particularly in dryland and semi-arid agricultural systems across West Asia, South Asia, and the Mediterranean Basin. Globally, lentil is produced on approximately 5.7 million hectares with a total output exceeding 7.1 million tons, with Canada, Australia, India, and Türkiye ranking among the leading producers [1]. Red lentil production is mostly concentrated in the Southeastern Anatolia Region of Türkiye, in which Şanlıurfa, Diyarbakır, and Siirt spearhead most of

the national output [2]. This region's Mediterranean-continental climate creates a growing environment in which year-to-year variability in seasonal precipitation is a dominant driver of lentil yield performance, thereby making it a particularly relevant target region for crop model evaluation under contrasting seasonal conditions.

Seasonal precipitation has shown a pronounced decline in southeastern Türkiye, where growing-season rainfall drops from a long-term average of approximately 600 mm to as low as 377 mm in 2020–2021 [2]. Precipitation has dropped by approximately 20% over the past 25 years, and significant further declines are foreseen in Türkiye's arid and semi-arid regions [3]. This increasing climate variability translates directly into year-to-year fluctuations in lentil yield and genotype performance. Thus, it creates a strong demand for reliable simulation tools that can reproduce crop behavior across contrasting seasonal conditions.

Process-based crop simulation models offer a mechanistic and integrative framework for quantifying genotype-by-environment interactions and support breeding and management decisions under variable growing conditions. The Decision Support System for Agrotechnology Transfer (DSSAT) is a comprehensive suite of process-based crop simulation models that have been continuously refined over several decades that establishes itself as a cornerstone in agricultural systems modelling [4]. The platform has been widely applied to pulse crops including chickpea, cowpea, and soybean to assess genotype performance, optimize management, and evaluate climate change impacts across diverse agro-ecological environments [5,6].

Despite the widespread application of DSSAT for major grain legumes, the availability of dedicated lentil model has been limited recently. The CSM-CROPGRO-Lentil model was formally incorporated into DSSAT v4.8.5 via adaption of the CSM-CROPGRO-faba bean model [5]. Jing et al. [7] reported that the model simulated development stages with an *RMSE* (Root means square error) of 4 days and aboveground biomass with $nRMSE \leq 23\%$, with seed yields generally well simulated across calibration and validation datasets ($nRMSE = 19\%$), though overestimation was noted under humid conditions in eastern Canada. The model's hierarchical parameterization structure provides flexibility for representing genetic diversity across large genotype panels. Cultivar-specific parameters are typically estimated using the Generalized Likelihood Uncertainty Estimation (GLUE) method, a Bayesian approach that uses Monte Carlo sampling from prior parameter distributions and a Gaussian likelihood function to determine the best-fitting coefficients based on observed end-of-season data [8].

A critical and often underappreciated constraint in CROPGRO model calibration is the requirement for non-stressed growing conditions. It is well established that if water, nutrients, or other stresses affect the experiments in calibration process, the estimate of cultivar coefficients may be poor. Thus, calibration treatments should include crops growing in a minimum-stress or stress-free environment, and the use of data from multiple years and sites generates more robust coefficients [9]. However, these methodological requirements are increasingly difficult to satisfy in practice. As climate variability intensifies across semi-arid growing regions, identifying two or more consecutive near-optimal seasons for multi-year calibration is becoming increasingly unlikely. The dataset used in the present study exemplifies this challenge precisely. The 2019–2020 growing season offered near-optimal precipitation and temperature conditions, while following seasons (2020–2021 and 2021–2022) were both characterized by severe moisture deficits during the critical flowering and seed-filling periods.

A further challenge in applying CROPGRO-Lentil for genotype evaluation lies in the model's parameterization structure as it relates to drought response. Genotype coefficients are divided into three hierarchical groups in DSSAT crop models. Species-level coefficients common to all cultivars, ecotype coefficients shared within groups of cultivars, and

cultivar-specific coefficients that differentiate individual genotypes [10]. Drought tolerance mechanisms such as root architecture, osmotic adjustment capacity, and stomatal regulation are not represented at the cultivar level in CROPGRO, meaning all genotypes share the same species-level water stress response functions regardless of their physiological drought tolerance characteristics.

The present study was designed to address these intersecting challenges by calibrating and validating the CSM-CROPGRO-Lentil model for 43 lentil genotypes (42 genotypes from ICARDA and one local cultivar) across three growing seasons (2019–2022) under semi-arid rainfed conditions in southeastern Türkiye. Building on previously published multi-year field evaluations [2] and PEG-induced drought tolerance screening [11] for the same genotypes, this study integrates three independent data sources to evaluate aggregate model performance and the model's capacity to differentiate genotypic responses under contrasting seasonal conditions. The specific objectives were: (i) to calibrate CROPGRO-Lentil cultivar and ecotype parameters for 43 lentil genotypes; (ii) to independently validate the model against two different seasons; (iii) to evaluate the model's capacity to reproduce the relative performance ranking of genotypes under drought conditions using Spearman rank correlation analysis; (iv) to assess whether model simulation bias in drought years is systematically associated with genotypic drought tolerance characteristics; and (v) to investigate whether simulation error is structured by genotypic characteristics (phenological maturity group and seed size) that influence the timing and duration of exposure to water deficit during critical reproductive stages.

2. Materials and Methods

2.1. Study Site

Field experiments were conducted at the Agricultural Area of the Siirt University, Siirt, Türkiye (37°58'04" N, 41°51'17" E) during three consecutive growing seasons. Data from 2019–2020 season were used for calibration, while measurements from 2020–2021 and 2021–2022 were used for validation. The climate of the study area is classified as semi-arid Mediterranean with cool wet winters and hot dry summers [2].

2.2. Climatological Data

Daily weather data including maximum and minimum temperature, solar radiation, and precipitation were obtained from the Türkiye meteorological station (37°58'04" N, 41°51'17" E, 588 m above sea level) for the 2019–2020, 2020–2021, and 2021–2022 seasons. The three seasons exhibited markedly different precipitation patterns, particularly during the critical flowering and seed-filling stages (Figure 1). The 2019–2020 calibration season received the highest total seasonal rainfall, with notably high precipitation in March (143.5 mm) and April (112.0 mm), coinciding with the vegetative-to-reproductive transition period. In contrast, both validation seasons experienced severe drought during April (11.0 mm in 2020–2021 and 13.8 mm in 2021–2022). Mean temperatures in April were also 3–4 °C higher in the validation years (16.0 °C and 14.7 °C, respectively) compared to the calibration year (11.7 °C). This indicates concurrent heat and drought stress during the reproductive stage. Long-term average (LTA) values for the 1938–2020 period were calculated from the same station dataset and used as a reference baseline for seasonal comparisons.

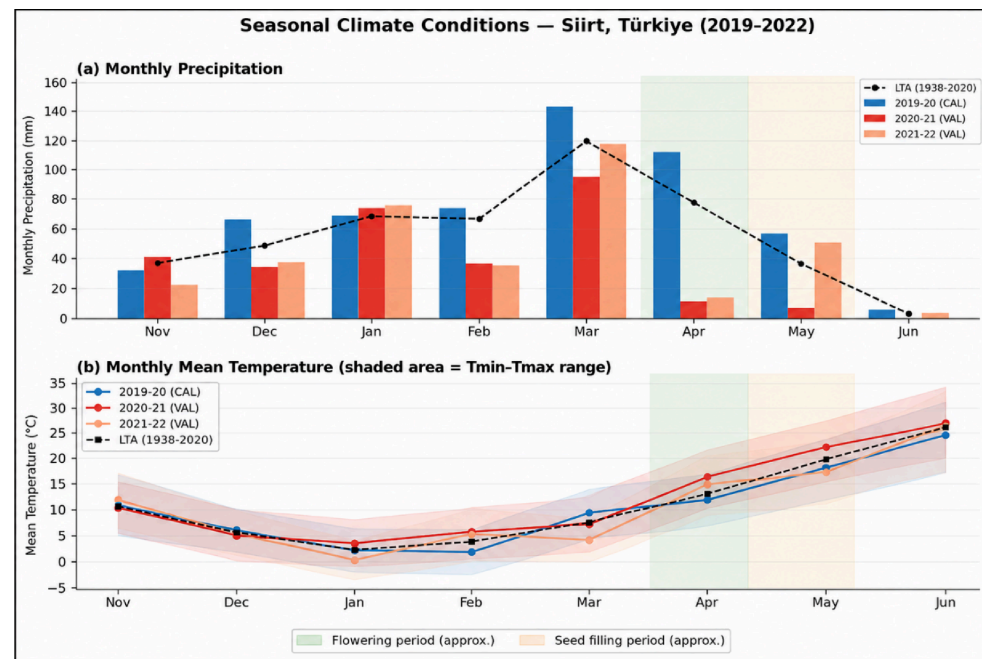


Figure 1. Seasonal distribution of monthly precipitation (a) and mean temperature (b) for the calibration (2019–2020) and validation (2020–2021, 2021–2022) growing seasons at Siirt, Türkiye. Shaded bars represent the approximate flowering and seed-filling periods based on observed phenological data. LTA = long-term average (1938–2020). Shaded regions indicate the approximate flowering period (April) and seed-filling period (May) based on mean observed anthesis and maturity dates across genotypes. Temperature shaded areas represent the daily minimum–maximum range for each season.

2.3. Soil Characterization

Soil physical and chemical properties were characterized for each growing season and entered DSSAT as season-specific soil profiles with a total depth of 150 cm divided into four layers (0–30, 30–60, 60–100, and 100–150 cm). Although all three experiments were conducted within the same research field, each growing season was established in a different plot area to avoid carry-over effects from the previous crop. Consequently, soil physical and chemical properties were characterized independently for each season, and the observed differences in soil texture across years (sandy clay loam in 2019–2020, silty loam in 2020–2021, and loam in 2021–2022) reflect genuine spatial variability within the field rather than measurement inconsistency. All profiles were characterized by slightly to moderately alkaline soil reaction, in which pH values increased with depth from 7.8 to 8.1, 7.3 to 7.6, and 8.0 to 8.3 across the three seasons, respectively. Soil albedo was set to 0.13 and the runoff curve number to 73 for all seasons because of the site's semi-arid dryland management conditions. Soil water retention parameters varied with texture. In the surface layer (0–30 cm), the lower limit, drained upper limit, and saturation ranged from 0.159–0.238, 0.266–0.344, and 0.374–0.459 $\text{cm}^3 \text{cm}^{-3}$ across seasons, respectively. Bulk density increased with depth in all seasons, thereby ranging from 1.35–1.47 g cm^{-3} in the surface layer to 1.52–1.58 g cm^{-3} in the deepest layer. Soil organic carbon was low across all profiles. It declined from 0.85–0.90% in the surface layer to 0.10–0.12% at 150 cm depth. Growth factor was set to 1.00 in the surface layer and decreased to 0.08 at 150 cm depth, which reflects the progressive restriction of rooting density with depth. Nutrient availability factors were set to 1.00 for all layers and seasons, thereby assuming non-limiting nitrogen and phosphorus conditions in the model. Soil characterization data for the field site across experimental seasons are also reported in Ceritoglu et al. [2].

2.4. Experimental Materials and Background

The plant material consisted of 42 lentil (*Lens culinaris* Medik.) accessions developed by ICARDA and one local check variety (Firat-87). All accessions were sourced from ICARDA's core collection (Lebanon) in 2016 and originated from two characterized germplasm groups: The Lentil International Elite Nursery-Extra Early was selected for short lifespan, while Lentil International Drought Tolerance Nursery was selected for high adaptability to drought conditions. Detailed pedigree information for all accessions and multi-year field performance data such as grain yield, biological yield, thousand-seed weight, and phenological traits are reported in Ceritoglu et al. [2], thus they are not repeated here.

A subset of 20 genotypes from this panel was previously screened for drought tolerance under controlled laboratory conditions using polyethylene glycol (PEG) 6000-induced osmotic stress at four concentration levels (0%, 10%, 15%, and 20%) during germination and early seedling stages [11]. According to the study results, fifteen germination and seedling traits were evaluated, and drought tolerance was quantified using the stress tolerance index (STI) computed for five traits including germination index (STI_GI), germination percentage (STI_GP), seedling vigor index (STI_SVI), dry weight (STI_SDW), and seedling length (STI_SL). Based on these STI values and principal component analysis, a composite STI was derived by normalizing and averaging all five STI parameters across the 20 genotypes, and genotypes were classified into tolerant ($n = 5$: G3664, G3840, G3829, G3828, G3710), intermediate ($n = 10$), and sensitive ($n = 5$: G3759, G3659, G3844, G3837, G35) groups using k-means clustering ($k = 3$) to enable statistically valid group comparisons in the present analysis.

2.5. Experimental Design

Experiments were conducted in a randomized complete block design with four replications. Sowing was performed on 21 November 2019, 15 November 2020, and 14 November 2021, respectively. Accessions were cultivated in 3 m² (1 m plot width and 3 m plot length) plots. Five rows were placed in each plot, and row spaces were arranged at 25 cm. The 250 seeds per square meter were manually sown, and the emergence rate exceeded 90% across all plots [2]. Region soils naturally harbor active *Rhizobium leguminosarum* strains, thereby promoting high nodulation on lentil roots. Nitrogen fertilizer and phosphorus were applied at 25 kg N ha⁻¹ and 60 kg P ha⁻¹ at sowing as a starter dose [12]. No supplemental irrigation was applied, i.e., crops were grown under rainfed conditions. Harvesting was carried out between the last week of May and the first week of June in all experimental seasons.

2.6. Experimental Data for DSSAT

Anthesis date (ADAT), maturity date (MDAT), thousand seed weight (TSW), grain yield and biological yield were used for calibration and validation processes. Thus, phenological and yield related characteristics were investigated in the model.

2.7. DSSAT-CROPGRO-Lentil Model

The DSSAT Cropping System Model (CSM) v.4.8.5 with the CROPGRO-Lentil module was used for all simulations. DSSAT simulates crop growth, development, and yield as a function of daily weather, soil water and nitrogen dynamics, and management inputs [4,13]. Water balance (RITCHIE method), soil evaporation (SULEIMAN-RITCHIE) and nitrogen fixation dynamic modules were activated. Carbon dioxide concentration was set at 410 ppm [14].

2.8. Model Calibration

Calibration was performed using data obtained from the 2019–2020 field experiment. Genotype-specific means of ADAT, MDAT, TSW, grain yield, and biological yield across the four field replications were used as calibration targets. These measurements were obtained from the field experiment described in Sections 2.5 and 2.6, and the underlying multi-year field data were previously reported by Ceritoglu et al. [2].

Three ecotype files were established to define ecotype-level parameters common to the ICARDA lentil accession group. Three ecotypes were established by applying k-means cluster analysis to 43 ICARDA lentil genotypes using three standardized variables (ADAT, MDAT and TSW) with the optimal cluster number ($k = 3$) independently confirmed by elbow method and the average silhouette width criterion (maximum silhouette width = 0.32 at $k = 3$). From each cluster, a pilot genotype was selected based on minimum Euclidean distance to the cluster centroid and used to calibrate the corresponding ecotype file (ECO1 ($n = 19$): intermediate-maturing and small-seeded; ECO2 ($n = 17$): late-maturing and medium-seeded; ECO3 ($n = 7$): intermediate-to-early maturing and large-seeded), after which the remaining genotypes within each cluster were individually calibrated using GLUE. This procedure ensured that phenological maturity and seed size were explicitly encoded at the ecotype level, providing a biologically grounded parameterization framework for the full 43-genotype panel.

Cultivar-level parameters were estimated using a two-stage procedure. In the first stage, the GLUE algorithm was applied to generate an optimal initial parameter set for each genotype through Monte Carlo sampling across pre-defined parameter ranges. The GLUE-derived parameter set was subsequently refined through a systematic iterative adjustment procedure, in which individual parameters were modified stepwise, and simulation output was compared against all observed variables until no further reduction in *RMSE* was achievable. This combined approach is consistent with established CROPGRO calibration practice, in which GLUE provides a statistically grounded starting solution that is then refined against the full suite of observed variables [8,15]. The calibrated coefficient values and definitions of the corresponding abbreviations for all genotypes are presented in Tables A1–A3 (Appendix A).

2.9. Model Validation

Independent validation was conducted using data from the 2020–2021 (Validation 1) and 2021–2022 (Validation 2) seasons. All genotype coefficients determined during calibration were kept unchanged. Separate experiment files were created for each validation year related to soil profiles and weather data.

2.10. Statistical Analysis and Model Performance Evaluation

Model performance during calibration and validation was assessed using four complementary statistical indices calculated for each output variable (ADAT, MDAT, TSW, grain yield, and biological yield). Statistical indices are Root mean square error (*RMSE*), normalized-*RMSE* (*nRMSE*), Nash–Sutcliffe efficiency (R^2), and Willmott's index of agreement (d) [16]. These indices were calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - X_i)^2}$$

$$nRMSE (\%) = \frac{RMSE}{\bar{X}} \times 100$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$d = 1 - \frac{\sum_{i=1}^n (Y_i - X_i)^2}{\sum_{i=1}^n (|Y_i - \bar{X}| + |X_i - \bar{X}|)^2}$$

where X_i and Y_i denote the observed and simulated values, respectively. $\bar{X}$ is the mean of observed values; and (n) is the number of observations.

Spearman rank correlation (*rho*) was computed between observed and simulated grain yield rankings across all 43 genotypes to evaluate the model's capacity to differentiate genotypic performance beyond aggregate error metrics. Spearman rank correlation was also used to examine the relationship between drought tolerance and grain yield bias in the two validation years. To enable statistically valid group comparisons, a composite STI was derived by normalizing and averaging all five STI parameters reported by Çiğ et al. [11] (STI_GL, STI_GP, STI_SVI, STI_SDW, STI_SL), and genotypes were re-classified into tolerant ($n = 5$), intermediate ($n = 10$), and sensitive ($n = 5$) groups using k-means clustering ($k = 3$). Differences in absolute bias between tolerant and sensitive groups were tested using the Mann–Whitney U-test.

Genotypes were classified into phenological groups based on observed anthesis dates in the calibration year (early: ADAT < 125 days; mid: 125–132 days; late: >132 days) and seed size groups based on observed TSW (small: <25 g; medium: 25–35 g; large: >35 g) to further investigate whether model simulation error was associated with genotypic characteristics. The calibration year was used as a phenological reference to reflect intrinsic genotypic characteristics under non-stressed conditions. Differences in grain yield bias among groups were tested using the Kruskal–Wallis test, followed by pairwise Wilcoxon rank-sum tests with Bonferroni correction where a significant group effect was detected ($p < 0.05$). The continuous relationship between TSW and grain yield bias was additionally quantified using Spearman rank correlation. All group-based analyses were performed separately for each validation year to account for the contrasting drought patterns of the two seasons. All statistical analyses were performed in R version 4.5.1 [17] using the *dplyr* [18], *tidyr* [19], *ggplot2* [20], and *ggrepel* [21] packages.

3. Results

3.1. Calibration Results of the CSM-CROPGRO-Lentil Model

Calibration performance was evaluated separately for phenological traits (ADAT and MDAT), thousand-seed weight (TSW), grain yield, and biological yield. Overall, the model achieved excellent to good simulation accuracy across all output variables, with performance varying markedly between phenological and yield-related traits (Table A4, Appendix B).

ADAT, MDAT and TSW simulations achieved the highest accuracy of all output variables with *nRMSE* of 0.8% (Figure 2a), 1.2% (Figure 2b) and 1.6% (Figure 2c), respectively. Grain yield simulation during calibration was classified as good with *RMSE* of 412.03 kg ha^{−1} and *nRMSE* of 19.3%. The correlation coefficient ($R^2 = 0.792$) and index of agreement ($d = 0.935$) indicated a strong positive relationship between observed and simulated yields (Figure 2d). Biological yield simulation similarly performed with *RMSE* = 833.75 kg ha^{−1}, *nRMSE* = 11.1%, $R^2 = 0.812$, $d = 0.941$, and a near-zero bias of −11.12 kg ha^{−1} (PBias = 1.3%), indicating balanced performance across the full range of biomass values (Figure 2e).

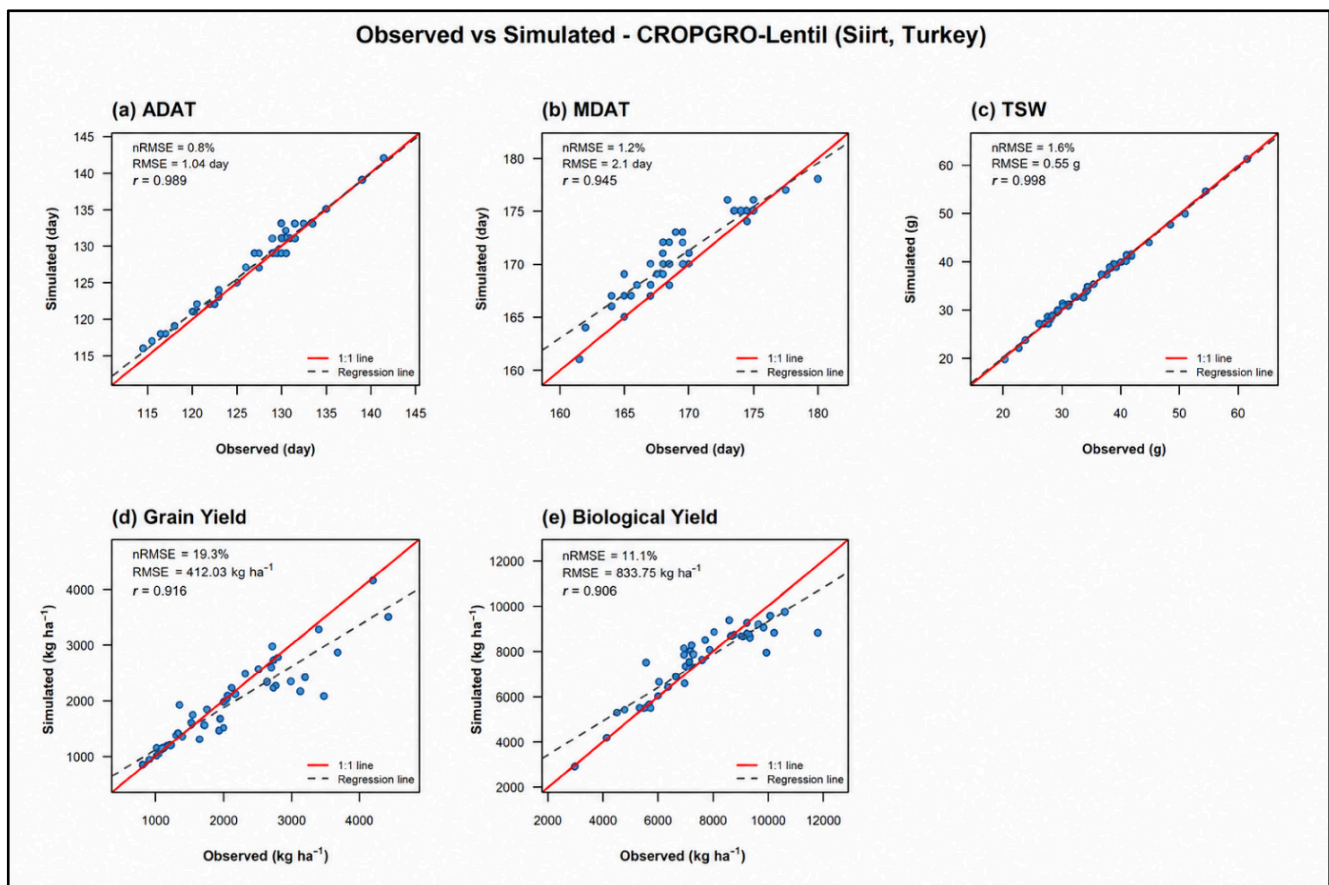


Figure 2. Observed versus simulated values for anthesis date (ADAT), maturity date (MDAT), thousand-seed weight (TSW), grain yield, and biological yield of 43 lentil genotypes during the calibration year (2019–2020). The red solid line represents the 1:1 line and the grey dashed line represent the linear regression line between observed and simulated values. TSW: Thousand-seed weight (g). $nRMSE$: Normalized root mean square error; $RMSE$: Root mean square error; r : Pearson correlation coefficient. $n = 43$ genotypes for all variables.

3.2. Validation Results of the CSM-CROPGRO-Lentil Model

3.2.1. 2020–2021 Validation Season

The 2020–2021 growing season represented a severe drought year, in which April precipitation reduced to 11.0 mm (approximately 10% of the calibration) and mean April temperature 4.3 °C higher than calibration season. Model performance markedly diverged between phenological and yield-related variables due to these conditions (Table A4, Appendix B).

ADAT simulation remained acceptable, although performance deteriorated relative to the calibration year. The model underestimated ADAT by approximately 8–9 days across the genotype panel (Bias = −8.55 days; PBias = −6.1%), resulting in an $RMSE$ of 9.47 days, $nRMSE$ of 6.8%, $R^2 = 0.635$, and $d = 0.603$ (Figure 3a). MDAT simulation showed a similar pattern with $RMSE = 8.45$ days, $nRMSE = 4.9\%$, $R^2 = 0.589$, $d = 0.315$, and a negative bias of −8.16 days (PBias = −4.7%) (Figure 3b).

TSW simulations deteriorated substantially under drought stress with $RMSE = 7.93$ g and $nRMSE = 22.9\%$. A negative bias of −7.02 g (PBias = −20%) indicated systematic underestimation of seed size across genotypes (Figure 3c).

In contrast, grain yield and biological yield simulations performed poorly during the 2020–2021 drought season. Grain yield $RMSE$ reached 525.07 kg ha⁻¹ with $nRMSE$ of 42.97%. The very weak association between observed and simulated grain yield values ($R^2 = 0.007$), together with the pronounced negative bias (−343.02 kg ha⁻¹;

PBias = -23.9%), demonstrated that the model systematically underestimated grain yield across the genotype panel (Figure 3d). Biological yield similarly exhibited a poor response with $RMSE = 1585.89 \text{ kg ha}^{-1}$, $nRMSE = 33.6\%$, $R^2 = 0.022$, $d = 0.318$, and a negative bias of $-910.79 \text{ kg ha}^{-1}$ (PBias = -15.6%) (Figure 3e). Consistent with these findings, the Spearman rank correlation between observed and simulated grain yield was non-significant ($\rho = -0.020$, $p = 0.897$), which confirms a loss of genotype-ranking capacity in the model's ability to differentiate genotypic performance under severe drought stress.

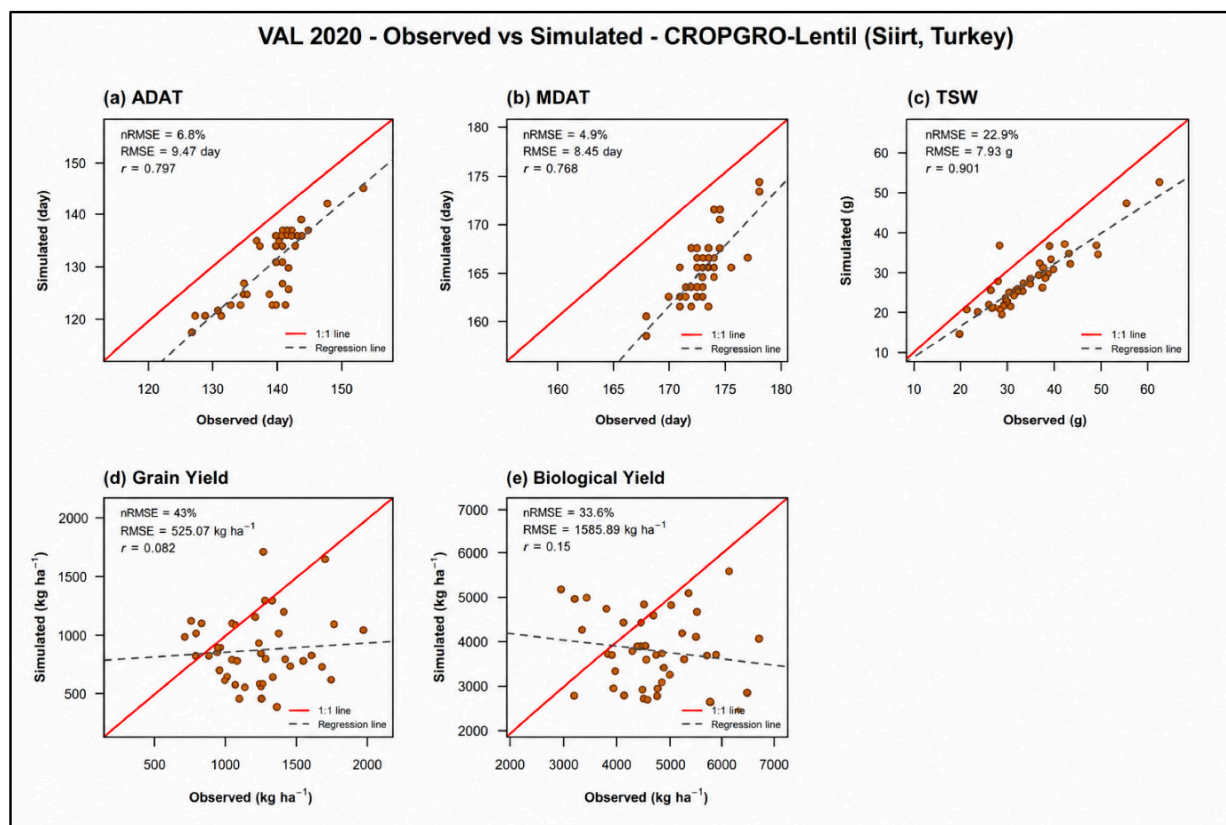


Figure 3. Observed versus simulated values for anthesis date (ADAT), maturity date (MDAT), thousand-seed weight (TSW), grain yield, and biological yield of 43 lentil genotypes during the first validation year (2020–2021). The red solid line represents the 1:1 line and the grey dashed line represent the linear regression line between observed and simulated values. TSW: Thousand-seed weight (g). $nRMSE$: Normalized root mean square error; $RMSE$: Root mean square error; r : Pearson correlation coefficient. $n = 43$ genotypes for all variables.

3.2.2. 2021–2022 Validation Season

The 2021–2022 growing season was also characterized by severely reduced April precipitation (13.8 mm), although May precipitation partially recovered to 50.4 mm. Mean observed grain yield declined further to 868 kg ha^{-1} , representing only 40.6% of the calibration-year mean and making this the most severely stressed season in the dataset (Table A4, Appendix B).

ADAT simulation recovered to near-calibration accuracy with $RMSE = 2.23$ days, $nRMSE = 1.6\%$, $R^2 = 0.647$, $d = 0.827$, and a small positive bias of 1.31 days (PBias = 0.9%) (Figure 4a). MDAT performance was comparatively weaker with $RMSE = 4.58$ days, $nRMSE = 2.6\%$, $R^2 = 0.308$, $d = 0.678$, and a negative bias of -2.24 days (PBias = -1.3%) (Figure 4b). TSW exhibited moderate performance with $RMSE = 7.06$ g, $nRMSE = 20.3\%$, $R^2 = 0.421$, $d = 0.788$, and a negligible bias (PBias = -0.9%) of -0.50 g (Figure 4c).

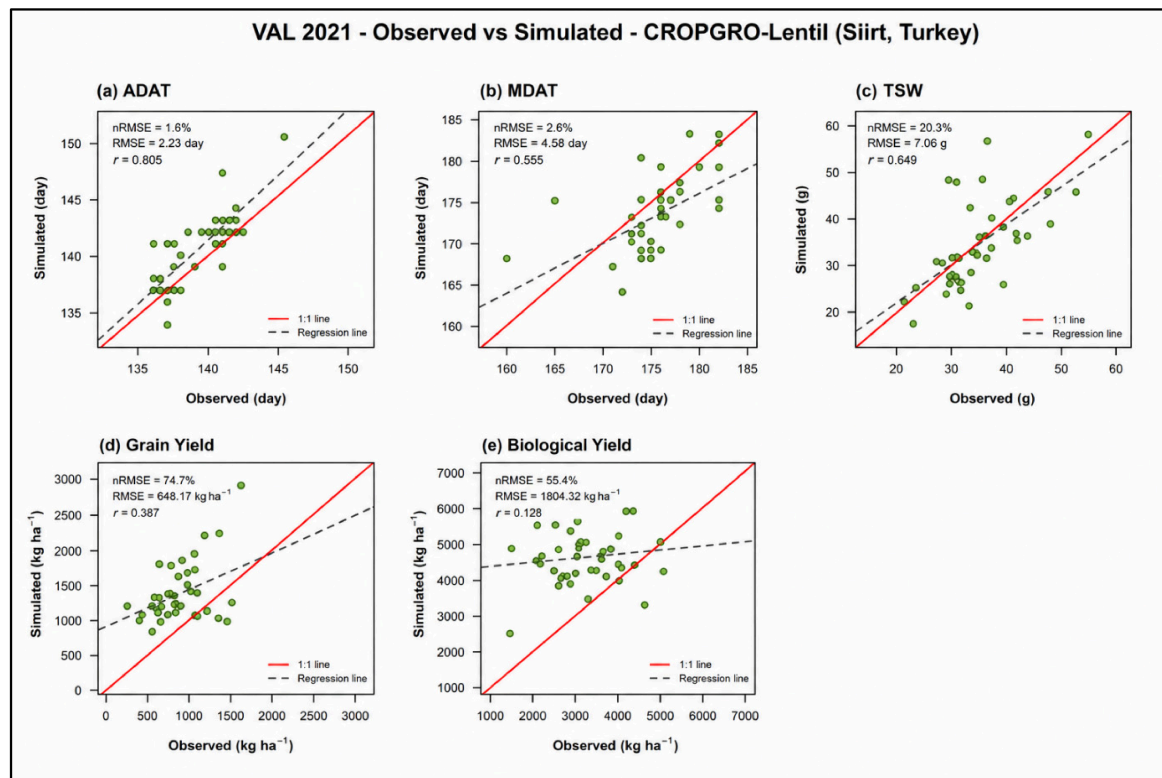


Figure 4. Observed versus simulated values for anthesis date (ADAT), maturity date (MDAT), thousand-seed weight (TSW), grain yield, and biological yield of 43 lentil genotypes during the second validation year (2021–2022). The red solid line represents the 1:1 line and the grey dashed line represent the linear regression line between observed and simulated values. TSW: Thousand-seed weight (g). *nRMSE*: Normalized root mean square error; *RMSE*: root mean square error; *r*: Pearson correlation coefficient. *n* = 43 genotypes for all variables.

In contrast, grain yield and biological yield simulations remained poor during the 2021–2022 validation season. Grain yield was substantially overestimated with $RMSE = 648.17 \text{ kg ha}^{-1}$, $nRMSE = 74.7\%$, $R^2 = 0.150$, $d = 0.453$, and a large positive bias of $+510.18 \text{ kg ha}^{-1}$ ($PBias = +75.4\%$) (Figure 4d). Biological yield similarly exhibited a poor response with $RMSE = 1804.32 \text{ kg ha}^{-1}$, $nRMSE = 55.4\%$, $R^2 = 0.016$, $d = 0.389$, and a positive bias of $+1446.45 \text{ kg ha}^{-1}$ ($PBias = +54.5\%$) (Figure 4e). Consistent with these findings, the Spearman rank correlation between observed and simulated grain yields remained non-significant ($\rho = +0.263$, $p = 0.088$), indicating that the model's ability to rank genotypes according to yield performance did not recover in the second validation year.

Taken together, the two validation years revealed a consistent structural pattern. The model retained moderate predictive capacity for phenological development across contrasting seasons, partially maintained its ability to simulate TSW, but failed to reproduce either the magnitude or the genotypic variation of grain and biological yield under drought-stressed conditions. Critically, the direction of yield bias differed between the two validation years with systematic underestimation observed in 2020–2021 and overestimation in 2021–2022. This pattern suggests that prediction errors were associated with the specific seasonal characteristics of drought rather than with a fixed directional bias in the calibrated parameter set.

3.3. Genotype-Specific Model Performance in Relation to Phenological and Seed Size Characteristics

The relationship between grain yield bias and two key genotypic characteristics (ADAT and TSW) was examined for both validation years. Genotypes were classified into phenological groups based on their observed ADAT in the stress-free calibration year

(2019): Early-maturing (ADAT < 125 days, $n = 15$), mid-maturing (125–132 days, $n = 22$), and late-maturing (>132 days, $n = 6$).

No significant difference in grain yield bias was observed among phenological groups in the 2020–2021 validation year (Kruskal–Wallis: $\chi^2 = 1.988$, $df = 2$, $p = 0.370$). In contrast, phenological group had a significant effect on simulation bias in the 2021–2022 season ($\chi^2 = 8.766$, $df = 2$, $p = 0.013$). Post hoc pairwise comparisons with Bonferroni correction revealed that late-maturing genotypes were simulated significantly more accurately than both early-maturing ($p = 0.043$) and mid-maturing genotypes ($p = 0.014$), while no significant difference was found between the early and mid-groups ($p = 1.000$). Mean grain yield bias for the late-maturing group in 2021–2022 was only +21.2 kg ha^{−1} compared to +525.4 and +632.9 kg ha^{−1} for the early and mid-groups, respectively (Figure S2, Supplementary File).

TSW was also significantly associated with simulation bias in the 2020–2021 validation year. Spearman rank correlation revealed a significant negative relationship between observed TSW and grain yield bias ($\rho = -0.410$, $p = 0.006$), indicating that large-seeded genotypes (TSW > 35 g) were more severely underestimated by the model (mean bias = −520.2 kg ha^{−1}) compared to medium-seeded (−260.1 kg ha^{−1}) and small-seeded genotypes (+17.8 kg ha^{−1}). This relationship was not significant ($\rho = -0.204$, $p = 0.190$) in the 2021–2022 season (Figure 5).

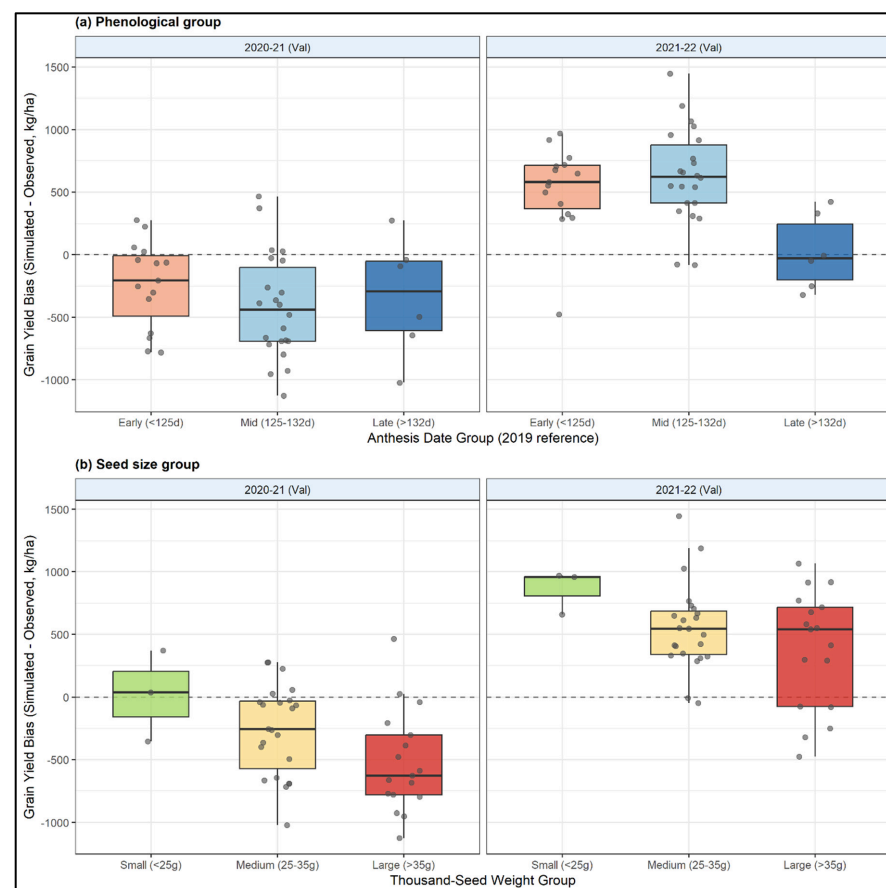


Figure 5. Grain yield bias (simulated – observed) of CROPGRO-Lentil across phenological groups (a) and seed size groups (b) for the 2020–2021 and 2021–2022 validation years. Phenological groups were defined based on observed anthesis dates in the stress-free calibration year (2019): early-maturing (ADAT < 125 days, $n = 15$), mid-maturing (125–132 days, $n = 22$), and late-maturing (>132 days, $n = 6$). Seed size groups were defined based on observed thousand-seed weight (TSW): small (<25 g), medium (25–35 g), and large (>35 g). The dashed horizontal line indicates zero bias. Points above the line indicate model overestimation; points below indicate underestimation.

4. Discussion

The calibration of CROPGRO-Lentil for 43 ICARDA lentil genotypes under semi-arid conditions in southeastern Türkiye yielded satisfactory results for phenological traits and thousand-seed weight with *nRMSE* values remaining below 2% for ADAT and MDAT, and below 1.6% for TSW in the calibration year. These results are consistent with previous CROPGRO-based calibration studies for grain legumes, which generally report reliable phenological simulations when cultivar parameters are adequately defined [7]. Grain yield and biological yield simulations were also acceptable during calibration (*nRMSE* = 19.3% and 11.1%, respectively) that confirm the suitability of the GLUE-based parameter estimation approach adopted in this study.

In both validation years, phenological simulations remained reliable with ADAT *nRMSE* values of 6.8% and 1.6% in 2020–2021 and 2021–2022, respectively. Although TSW simulations retained moderate agreement in the validation years, their accuracy declined substantially relative to the calibration season, that indicate seed-size simulation was also affected by drought conditions. These findings are consistent with previous process-based crop modelling studies demonstrating that phenological and morphological traits are generally more stable across environments than yield-related traits [6,7].

However, grain yield and biological yield simulations deteriorated substantially in both validation years. The model systematically underestimated grain yield in 2020–2021 (*nRMSE* = 43.0%, Bias = -343 kg ha^{-1}) and overestimated it in 2021–2022 (*nRMSE* = 74.7%, Bias = $+510 \text{ kg ha}^{-1}$), revealing an inconsistent bias direction across years. This pattern is consistent with CROPGRO applications under semi-arid conditions reported by Suleiman et al. [6], who found that CROPGRO-Soybean failed to satisfactorily simulate yield, biomass, and TSW under water stress imposed during reproductive stages, and concluded that the model could not adequately track observed yield patterns under such conditions. Inspection of simulated water stress factors (SWFAC) extracted from the DSSAT Overview output files for a representative subset of seven genotypes revealed that soil hydraulic properties contributed to the contrasting bias directions observed between validation seasons. During the critical seed-filling period (first seed to physiological maturity), mean growth-limiting SWFAC across the seven genotypes was 0.274 in 2020–2021 and 0.155 in 2021–2022, indicating that the model computed substantially higher water stress intensity in the first validation season. This difference is mechanistically consistent with the contrasting soil profiles: the silty loam profile used in 2020–2021 had a lower drained upper limit (DUL = $0.266 \text{ cm}^3 \text{ cm}^{-3}$) than the loam profile in 2021–2022 (DUL = $0.344 \text{ cm}^3 \text{ cm}^{-3}$), meaning that equivalent rainfall inputs resulted in lower plant-available water in the first validation season. The combination of lower DUL and the near-total April precipitation deficit (11.0 mm) would have triggered earlier and more severe stress onset in the DSSAT water balance module in 2020–2021, contributing to the systematic underestimation observed in that season. Conversely, the higher water-holding capacity of the 2021–2022 loam profile, combined with the partial May precipitation recovery (50.4 mm), may have allowed the model to maintain higher simulated assimilate supply than the crop achieved under field conditions, contributing to overestimation. These results suggest that inter-annual soil profile variability partially confounded the atmospheric drought signal in both validation seasons. Disentangling these contributions conclusively would require a controlled scenario analysis using a standardized soil profile across all three seasons—an investigation recommended for future multi-site validation studies.

Importantly, the observed inter-annual decline in mean grain yield—from 2136 kg ha^{-1} in 2019–2020 to 1222 kg ha^{-1} in 2020–2021 and 868 kg ha^{-1} in 2021–2022—was broadly consistent with the yield trajectory reported by Ceritoglu et al. [2] for the same genotype set under the same field conditions, confirming that the model correctly captured the

direction of seasonal yield response. However, the critical failure was not in tracking the seasonal trend but in reproducing the relative performance of individual genotypes within each year. Spearman rank correlation between observed and simulated grain yields across all 43 genotypes was highly significant during calibration ($\rho = +0.930$, $p < 0.001$) but collapsed to non-significant levels in both validation years (2020–2021: $\rho = -0.020$, $p = 0.897$; 2021–2022: $\rho = +0.263$, $p = 0.088$), indicating a complete loss of genotypic differentiation capacity under drought stress conditions. At the individual genotype level, only three of 43 accessions (G37, G3695, and G21151) maintained grain yield simulation error below 40% across both drought-affected validation seasons, further illustrating the severity of this breakdown.

This combination of terminal drought and elevated temperature during reproductive development is known to cause disproportionate yield losses in lentils that are difficult to simulate with process-based models [22,23]. The DSSAT water stress module computes drought stress factors (SWFAC) based on species-level soil water threshold parameters defined in the SPE file, applying identical stress response functions to all genotypes regardless of their inherent drought tolerance mechanisms. This structural characteristic of the model likely contributed to its inability to differentiate genotypic responses under the observed terminal drought conditions. The limitations of default water stress response functions in DSSAT have been previously highlighted. Yao et al. [24] demonstrated that modifying these functions in DSSAT-CERES-Wheat improved overall simulation accuracy by 13.7% under water deficit conditions, particularly for biomass and grain yield, underscoring the sensitivity of DSSAT model outputs to the formulation of water stress response functions at the species level.

To further investigate whether the model's failure was linked to genotypic drought tolerance characteristics, grain yield bias in validation years was compared against stress tolerance indices (STI) derived from a PEG-induced drought stress experiment conducted on 19 panel genotypes [11]. Mann–Whitney U-tests remained non-significant in both validation seasons (2020–2021: $U = 12.0$, $p = 1.000$; 2021–2022: $U = 16.0$, $p = 0.548$), and Spearman correlation between composite STI scores and absolute grain yield bias was also non-significant across all 20 genotypes with PEG data (2020–2021: $\rho = -0.081$, $p = 0.733$; 2021–2022: $\rho = +0.093$, $p = 0.698$). Notably, in the 2021–2022 season, the tolerant group showed a higher mean absolute bias (726 kg ha^{-1}) than the sensitive group (590 kg ha^{-1}), further excluding any systematic relationship between PEG-based drought tolerance and model simulation accuracy. These results confirm that the absence of association between drought tolerance classification and simulation error is not an artefact of insufficient statistical power, but reflects a genuine structural characteristic of the model—its species-level water stress response functions do not accommodate genotype-specific variation in drought tolerance mechanisms such as root architecture, osmotic adjustment, or stomatal regulation, traits that were found to differ significantly among genotypes in the PEG experiment [11] and in field evaluations [2]. Similar structural constraints have been identified in other DSSAT applications, where modifying the water stress response functions at the species level substantially improved simulation accuracy under drought conditions [24].

The finding that late-maturing genotypes (ADAT > 132 days in the calibration year) showed significantly lower grain yield bias in the 2021–2022 validation season (mean bias = $+21.2 \text{ kg ha}^{-1}$) compared to early and mid-maturing groups ($+525.4$ and $+632.9 \text{ kg ha}^{-1}$, respectively (Kruskal–Wallis $p = 0.013$ warrants careful interpretation). Figure S2 demonstrates that all three phenological groups reached anthesis within the April drought window in the 2021–2022 season, when precipitation totaled only 13.8 mm, confirming that the lower simulation bias observed for late-maturing genotypes cannot be attributed to phenological escape from the critical drought period [25,26]. Two alternative

explanations merit consideration. First, the partial precipitation recovery in May (50.4 mm) coincided with the post-anthesis seed-filling period of late-maturing genotypes, potentially providing a more favorable carbon assimilation environment that the model was better equipped to simulate compared to the near-complete moisture deficit experienced by earlier-maturing genotypes during their seed-filling stage. Second, it is possible that the cultivar-level parameters of the late-maturing group were more accurately estimated during calibration, conferring an artificial simulation advantage unrelated to their physiological response under drought. The late-maturing group comprised only six genotypes with a relatively narrow ADAT range, and the current data do not allow these two explanations to be distinguished. This finding is therefore best regarded as an unexplained observation that warrants investigation in future studies with larger genotype panels and replicated stress environments (Figure S2, Supplementary File).

The negative relationship between seed size and grain yield bias observed in 2020–2021 ($\rho = -0.410$, $p = 0.006$) adds a complementary dimension to this analysis. Large-seeded genotypes, which require sustained carbon assimilation and assimilate partitioning during an extended seed-filling period to achieve their potential seed size, were disproportionately underestimated by the model under the prolonged drought conditions of 2020–2021. This suggests that the model's water stress response functions, which reduce photosynthesis and assimilate supply uniformly across all genotypes once soil water falls below species-level threshold values, may not adequately capture the differential sensitivity of seed-filling dynamics in genotypes with contrasting seed size potential. Furthermore, post-anthesis drought stress accelerates the remobilization of pre-anthesis reserves stored in vegetative tissues, yet the contribution of this remobilization to grain filling increases substantially under prolonged drought stress compared to well-watered conditions [27], and this process is not explicitly represented in CROPGRO's carbon balance routines [4]. The absence of a significant TSW–bias relationship in 2021–2022 further supports this interpretation: in that season, the drought was concentrated in a brief terminal window rather than sustained across the seed-filling period, and the differential demand for prolonged carbon supply between large- and small-seeded genotypes was less consequential.

These results indicate that future studies could examine whether changes to CROPGRO-Lentil's representation of genotype-specific drought responses improve its performance under water-limited conditions. Sensitivity analyses could be used to investigate the potential contribution of parameters related to root water uptake (RWUEP1) and soil water stress response thresholds (XSWFAC and YSWFAC). Their possible relationships with measurable physiological indicators, such as root biomass, seedling vigour index, or osmotic potential under stress, could also be explored [11]. Because these possibilities were not tested in the present study, their effectiveness in improving the predictive performance of CROPGRO-Lentil remains to be experimentally evaluated. Future research could similarly assess whether a more explicit representation of pre-anthesis carbon remobilization and phenology-dependent stress sensitivity improves genotype-specific simulations under variable drought regimes.

An important limitation of this study is that cultivar coefficients were estimated using a single minimum-stress season. Although this approach reduced the confounding influence of drought during calibration, it may also have resulted in coefficients that were overly fitted to the environmental conditions of 2019–2020. Consequently, the poor yield predictions in the validation years cannot be attributed exclusively to structural limitations in the model's drought-response functions. Insufficient representation of environmental variability during calibration, uncertainty in cultivar coefficient estimation, and interannual differences in soil hydraulic properties may also have contributed to the observed errors. The present dataset does not allow the relative contributions of these factors to be separated.

5. Conclusions

This study demonstrated that the CSM-CROPGRO-Lentil model, when parameterized through a hierarchical GLUE-based approach for a large and genetically diverse panel of 43 ICARDA lentil genotypes, achieves reliable simulation of phenological development and seed size under near-optimal growing conditions. The calibration results (anthesis date $nRMSE$ below 1% and thousand-seed weight $nRMSE$ below 2%) indicate that cultivar-level parameterization within a three-ecotype framework captured much of the phenological and morphological variation within the calibration dataset. Grain yield calibration performance ($nRMSE = 19.3\%$) was consistent with previously reported CROPGRO-Lentil benchmarks, and genotype ranking capacity was high ($\rho = +0.930$).

However, validation under two terminal-drought seasons revealed a substantial decline in model transferability, particularly for grain yield, biological yield, and genotype ranking. This decline may reflect the interacting effects of limitations in the model's drought-response functions, potential overfitting to the single calibration season, uncertainty in cultivar coefficient estimation, and interannual differences in soil hydraulic properties. The present study cannot separate the relative contributions of these factors. At the individual genotype level, only three accessions (G37, G3695, and G21151) maintained acceptable simulation accuracy ($APE < 40\%$) across both drought-affected validation seasons, underscoring the severity of model failure under terminal water deficit.

These findings carry practical implications for the use of CROPGRO-Lentil in genotype evaluation and breeding support under semi-arid conditions. The model is suitable for simulating phenological adaptation and seasonal yield trends but should not be used to rank genotypes by yielding potential under drought-stressed seasons without substantial modifications to its stress parameterization framework. Future research should investigate whether modifications to the model's representation of genotype-specific drought responses could improve its predictive performance under water-limited conditions. Sensitivity analyses could be conducted to evaluate the potential contribution of parameters related to root water uptake (RWUEP1) and soil water stress response thresholds (XSWFAC and YSWFAC). Their possible relationships with independently measured physiological traits or stress tolerance indicators could also be explored. Because the predictive effectiveness of these parameters was not tested in the present study, their usefulness for improving CROPGRO-Lentil performance remains to be experimentally evaluated.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/agronomy16151441/s1>, Figure S1: Observed versus simulated values for anthesis date (a), maturity date (b), thousand-seed weight (c), grain yield (d), and biological yield (e) of 43 lentil genotypes across the calibration year (2019, Cal) and two validation years (2020, Val; 2021, Val) at Siirt, Türkiye. Figure S2: Observed anthesis date (ADAT) of 43 lentil genotypes grouped by phenological maturity class across the three experimental seasons.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available in this article and in the Supplementary Materials.

Acknowledgments: The author would like to thank Figen Ceritoglu for her valuable support in statistical analyses. During the preparation of this manuscript, the author used Claude (Anthropic, Sonnet-5, San Francisco, CA, USA) and ChatGPT (OpenAI, GPT-4, San Francisco, CA, USA) for the purposes of literature searching and text editing/proofreading only. These tools were not used for interpretation or discussion of results, nor for directly generating any substantive content of this publication. The author has reviewed and edited all outputs and takes full responsibility for the content of this publication.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADAT	Anthesis date
CSM	Crop Simulation model
DSSAT	Decision Support System for Agrotechnology Transfer
DUL	Drained upper limit
GLUE	Generalized Likelihood Uncertainty Estimation
ICARDA	International Center for Agricultural Research in the Dry Areas
MDAT	Maturity date
nRMSE	Normalized Root Mean Square Error
RMSE	Root Mean Square Error
STI	Stress tolerance index
TSW	Thousand seed weight

Appendix A

Table A1. Calibrated phenological cultivar coefficients for 43 lentil genotypes used in the CSM-CROPGRO-Lentil model.

Genotype	ECO	CLSD	PPSEN	EM-FL	FL-SH	FL-SD	SD-PM	FL-LF
G3737	LAT001	15.80	0	6.8	3.8	8.0	13.30	20.00
G3710	LAT001	15.80	−0.11	7.2	5.2	9.5	12.90	20.00
G3839	LAT001	15.80	−0.11	8.4	5.0	9.0	10.20	20.00
G3828	LAT001	15.80	−0.11	8.7	5.0	9.0	11.00	20.00
G3690	LAT001	15.80	−0.11	7.5	5.0	9.0	11.60	20.00
G3664	LAT001	15.80	−0.11	8.5	5.0	9.0	11.60	20.00
G3695	LAT001	15.80	−0.11	7.0	5.0	9.0	11.60	20.00
G87	LAT001	15.80	−0.11	6.5	5.0	9.0	11.30	20.00
G3716	LAT001	15.80	−0.11	6.5	5.0	9.0	11.00	20.00
G3659	LAT001	15.80	−0.11	8.5	5.0	9.0	11.90	20.00
G3703	LAT001	15.80	−0.11	8.1	5.0	9.0	11.60	20.00
G3829	LAT001	15.80	−0.11	8.4	5.0	9.0	11.40	20.00
G3687	LAT001	15.80	−0.11	6.8	7.3	9.9	11.40	20.00
G3743	LAT001	15.80	−0.11	6.3	7.3	9.9	11.40	20.00
G3701	LAT001	15.80	−0.11	8.5	7.5	10.4	8.00	20.00
G3705	LAT001	15.80	−0.11	7.4	7.5	8.4	12.50	20.00
G3715	LAT001	15.80	−0.11	6.0	7.5	8.4	10.80	20.00
G3689	LAT001	15.80	−0.11	8.4	7.9	8.9	10.80	20.00
G3696	LAT001	15.80	−0.11	8.4	7.9	8.9	10.80	20.00
G3679	LAT002	15.80	−0.11	9.0	5.2	9.5	10.50	20.00
G21151	LAT002	15.80	−0.11	10.5	5.2	9.5	13.40	20.00
G37	LAT002	15.80	−0.11	11.5	5.2	9.5	11.00	20.00
G3805	LAT002	15.80	−0.11	9.0	5.2	9.5	11.10	20.00
G3649	LAT002	15.80	−0.11	8.0	5.2	9.5	12.60	20.00
G3697	LAT002	15.80	−0.11	9.0	5.2	9.5	11.10	20.00
G3840	LAT002	15.80	−0.11	9.5	5.2	9.5	13.30	20.00
G3678	LAT002	15.80	−0.11	9.0	5.2	9.5	12.80	20.00
G3658	LAT002	15.80	−0.11	8.8	5.2	9.5	13.60	20.00
G3673	LAT002	15.80	−0.11	8.8	5.2	9.5	16.20	20.00
G3662	LAT002	15.80	−0.11	8.8	5.6	9.7	11.90	20.00
G3713	LAT002	15.80	−0.11	8.1	5.2	9.5	15.00	20.00

Table A1. Cont.

Genotype	ECO	CLSD	PPSEN	EM-FL	FL-SH	FL-SD	SD-PM	FL-LF
G3726	LAT002	15.80	−0.11	8.7	5.2	9.5	15.00	20.00
G3837	LAT002	15.80	−0.11	8.9	5.2	9.5	12.00	20.00
G3819	LAT002	15.80	−0.11	8.7	5.2	9.5	13.00	20.00
G3653	LAT002	15.80	−0.11	8.5	5.2	9.5	14.00	20.00
G3844	LAT002	15.80	−0.11	9.0	5.2	9.5	11.50	20.00
G35	LAT003	15.80	−0.11	6.8	3.5	7.0	14.40	20.00
G3744	LAT003	15.80	−0.11	6.8	5.2	8.9	12.80	20.00
G3759	LAT003	15.80	−0.11	7.0	5.2	9.0	12.50	20.00
G3750	LAT003	15.80	−0.11	6.6	6.2	9.9	13.00	20.00
G3780	LAT003	15.80	−0.11	7.0	6.0	9.0	10.00	20.00
G3761	LAT003	15.80	−0.11	6.8	6.0	9.0	13.90	20.00
G3771	LAT003	15.80	−0.11	8.8	7.1	10.3	11.10	20.00

ECO: Ecotype; CSDVAR: critical short-day length (h); PPSEN: slope of the relative response of development to photoperiod with time (h^{-1}); EM-FL: thermal time from emergence to first flower (photothermal days); FL-SH: thermal time from first flower to first pod (photothermal days); FL-SD: thermal time from first flower to first seed (photothermal days); SD-PM: thermal time from first seed to physiological maturity (photothermal days); FL-LF: thermal time from first flower to the end of leaf expansion (photothermal days).

Table A2. Calibrated growth and morphological cultivar coefficients for 43 lentil genotypes used in the CSM-CROPGRO-Lentil model.

Genotype	ECO	LFMAX	SLAVR	SIZLF
G3737	LAT001	1.59	130	170
G3710	LAT001	1.92	299	326
G3839	LAT001	1.68	150	180
G3828	LAT001	1.87	140	168
G3690	LAT001	1.85	240	250
G3664	LAT001	1.75	240	250
G3695	LAT001	1.85	340	350
G87	LAT001	1.60	150	180
G3716	LAT001	1.90	290	365
G3659	LAT001	1.15	190	265
G3703	LAT001	1.90	310	385
G3829	LAT001	1.95	350	415
G3687	LAT001	2.15	350	415
G3743	LAT001	1.79	130	156
G3701	LAT001	1.99	350	420
G3705	LAT001	1.70	350	420
G3715	LAT001	1.80	350	420
G3689	LAT001	1.95	350	420
G3696	LAT001	1.80	350	420
G3679	LAT002	1.93	400	480
G21151	LAT002	1.33	100	120
G37	LAT002	0.92	140	168
G3805	LAT002	1.37	200	240
G3649	LAT002	1.30	100	120
G3697	LAT002	1.90	300	360
G3840	LAT002	1.95	300	360
G3678	LAT002	1.33	100	130
G3658	LAT002	2.25	470	490
G3673	LAT002	1.63	120	144
G3662	LAT002	1.99	400	480
G3713	LAT002	1.22	55	66

Table A2. Cont.

Genotype	ECO	LFMAX	SLAVR	SIZLF
G3726	LAT002	1.97	299	326
G3837	LAT002	1.99	400	480
G3819	LAT002	1.75	299	358
G3653	LAT002	1.91	350	426
G3844	LAT002	1.95	299	326
G35	LAT003	1.70	300	360
G3744	LAT003	1.99	347	428
G3759	LAT003	1.70	150	180
G3750	LAT003	1.60	200	240
G3780	LAT003	1.58	100	120
G3761	LAT003	1.30	100	124
G3771	LAT003	1.90	400	480

ECO: ecotype code; LFMAX: maximum leaf photosynthesis rate at 30 °C, 350 ppm CO₂, and high light (mg CO₂ m⁻² s⁻¹); SLAVR: specific leaf area under standard growth conditions (cm² g⁻¹); SIZLF: maximum size of a fully expanded leaf (cm²).

Table A3. Calibrated reproductive and seed-related cultivar coefficients for 43 lentil genotypes used in the CSM-CROPGRO-Lentil model.

Genotype	ECO	XFRT	WTPSD	SFDUR	SDPDV	PODUR	THRSH	SDPRO	SDLIP
G3737	LAT001	0.83	0.063	20.0	2.20	14.0	92.0	0.212	0.020
G3710	LAT001	0.95	0.046	20.0	2.20	14.0	94.0	0.245	0.020
G3839	LAT001	0.82	0.062	20.0	2.20	14.0	92.0	0.180	0.020
G3828	LAT001	0.60	0.064	20.0	2.20	14.0	90.0	0.160	0.020
G3690	LAT001	0.95	0.066	20.0	2.20	14.0	94.0	0.185	0.020
G3664	LAT001	0.95	0.076	20.0	2.20	14.0	94.0	0.280	0.020
G3695	LAT001	0.98	0.067	20.0	2.20	14.0	94.0	0.280	0.020
G87	LAT001	0.94	0.073	20.0	2.20	14.0	92.0	0.180	0.020
G3716	LAT001	0.97	0.083	20.0	2.20	14.0	95.0	0.320	0.020
G3659	LAT001	0.96	0.074	20.0	2.20	14.0	94.0	0.230	0.020
G3703	LAT001	0.80	0.070	20.0	2.20	14.0	94.0	0.180	0.020
G3829	LAT001	0.90	0.082	20.0	2.20	14.0	94.0	0.270	0.020
G3687	LAT001	0.97	0.070	20.0	2.20	14.0	96.0	0.300	0.020
G3743	LAT001	0.89	0.072	20.0	2.20	14.0	96.0	0.180	0.020
G3701	LAT001	0.97	0.105	20.0	2.20	14.0	95.0	0.290	0.020
G3705	LAT001	0.95	0.057	20.0	2.20	14.0	95.0	0.260	0.020
G3715	LAT001	0.97	0.073	20.0	2.20	14.0	95.0	0.300	0.020
G3689	LAT001	0.95	0.073	20.0	2.20	14.0	95.0	0.300	0.020
G3696	LAT001	0.97	0.077	20.0	2.20	14.0	96.0	0.350	0.020
G3679	LAT002	0.84	0.090	20.0	2.20	14.0	90.0	0.250	0.020
G21151	LAT002	0.57	0.077	20.0	2.20	14.0	92.0	0.180	0.020
G37	LAT002	0.97	0.095	23.0	2.20	14.0	94.0	0.300	0.020
G3805	LAT002	0.78	0.082	20.0	2.20	14.0	95.0	0.180	0.020
G3649	LAT002	0.75	0.058	20.0	2.20	14.0	92.0	0.180	0.020
G3697	LAT002	0.95	0.096	20.0	2.20	14.0	97.0	0.400	0.020
G3840	LAT002	0.60	0.060	20.0	2.20	14.0	92.0	0.180	0.020
G3678	LAT002	0.68	0.064	20.0	2.20	14.0	93.0	0.130	0.020
G3658	LAT002	0.95	0.076	20.0	2.20	14.0	97.0	0.370	0.020
G3673	LAT002	0.95	0.077	20.0	2.20	14.0	96.0	0.220	0.020
G3662	LAT002	0.90	0.090	20.0	2.20	14.0	94.0	0.200	0.020
G3713	LAT002	0.90	0.076	20.0	2.20	14.0	91.0	0.180	0.020
G3726	LAT002	0.62	0.079	20.0	2.20	14.0	93.0	0.235	0.020

Table A3. Cont.

Genotype	ECO	XFRT	WTPSD	SFDUR	SDPDV	PODUR	THRSH	SDPRO	SDLIP
G3837	LAT002	0.95	0.099	20.0	2.20	14.0	93.0	0.275	0.020
G3819	LAT002	0.95	0.082	20.0	2.20	14.0	93.0	0.245	0.020
G3653	LAT002	0.90	0.055	20.0	2.20	14.0	92.0	0.245	0.020
G3844	LAT002	0.90	0.094	20.0	2.20	14.0	94.0	0.245	0.020
G35	LAT003	0.97	0.098	20.0	2.20	14.0	98.6	0.270	0.020
G3744	LAT003	0.95	0.090	20.0	2.20	14.0	97.0	0.320	0.020
G3759	LAT003	0.95	0.099	20.0	2.20	14.0	93.0	0.197	0.020
G3750	LAT003	0.70	0.103	20.0	2.20	14.0	92.0	0.220	0.020
G3780	LAT003	0.87	0.159	20.0	2.20	14.0	92.0	0.185	0.020
G3761	LAT003	0.68	0.082	20.0	2.20	14.0	92.0	0.170	0.020
G3771	LAT003	0.97	0.132	20.0	2.20	14.0	98.0	0.450	0.020

ECO: ecotype code; XFRT: maximum fraction of daily growth partitioned to seed and shell; WTPSD: maximum weight per seed (g); SFDUR: seed-filling duration for a pod cohort under standard growth conditions (photothermal days); SDPDV: average number of seeds per pod under standard growing conditions; PODUR: time required to reach the final pod load under optimal conditions (photothermal days); THRSH: threshing percentage, defined as the ratio of seed weight to the combined pod and seed weight at maturity (%); SDPRO: fraction of protein in seed (g g^{-1}); SDLIP: fraction of lipid in seed (g g^{-1}).

Appendix B

Table A4. Statistical indicators of model performance during calibration and validation years.

	Trait	Unit	RMSE	nRMSE	R ²	d	Bias
2019–2020 Calibration	ADAT	day	1.04	0.8	0.971	0.993	0.5
	MDAT	day	2.1	1.2	0.784	0.939	1.45
	TSW	g	0.55	1.6	0.996	0.999	0.04
	Grain yield	kg ha^{-1}	412.03	19.3	0.792	0.935	−171.42
	Biological yield	kg ha^{-1}	833.75	11.1	0.812	0.941	−11.12
2020–2021 Validation	ADAT	day	9.47	6.793	0.6350	0.603	−8.55
	MDAT	day	8.45	4.889	0.5890	0.315	−8.16
	TSW	g	7.93	22.89	0.8110	0.793	−7.02
	Grain yield	kg ha^{-1}	525.07	42.97	0.0067	0.391	−343.02
	Biological yield	kg ha^{-1}	1585.89	33.62	0.0224	0.318	−910.79
2021–2022 Validation	ADAT	day	2.23	1.600	0.647	0.827	1.31
	MDAT	day	4.58	2.613	0.308	0.678	−2.24
	TSW	g	7.06	20.27	0.421	0.788	−0.5
	Grain yield	kg ha^{-1}	648.17	74.70	0.150	0.453	510.18
	Biological yield	kg ha^{-1}	1804.32	55.42	0.016	0.389	1446.45

(RMSE: Root Mean Square Error; nRMSE: Normalized Root Mean Square Error; R²: Coefficient of Determination; d: Willmott's Index of Agreement; Bias: Mean Bias Error).

References

1. FAO. FAOSTAT—Crop Production Data; Food and Agriculture Organization of the United Nations: Rome, Italy, 2025; Available online: <https://www.fao.org/faostat> (accessed on 10 June 2025).
2. Ceritoglu, M.; Çiğ, F.; Erman, M.; Ceritoglu, F. Integrating multi-trait selection indices for climate-resilient lentils: A three-year evaluation of earliness and yield stability under semi-arid conditions. *Agronomy* **2025**, *15*, 1554. [\[CrossRef\]](#)
3. FAO. Drought Trends and Impacts on Agricultural Production and Natural Resources in the Context of Climate Change; Food and Agriculture Organization of the United Nations: Rome, Italy, 2015.
4. 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. The DSSAT cropping system model. *Eur. J. Agron.* **2003**, *18*, 235–265. [\[CrossRef\]](#)
5. Boote, K.J.; Mínguez, M.I.; Sau, F. Adapting the CROPGRO legume model to simulate growth of faba bean. *Agron. J.* **2002**, *94*, 743–756. [\[CrossRef\]](#)
6. Suleiman, A.A.; Tojo Soler, C.M.; Hoogenboom, G. Deficit irrigations during soybean reproductive stages and CROPGRO-soybean simulations under semi-arid climatic conditions. *Field Crops Res.* **2007**, *101*, 140–149.

7. Jing, Q.; Boote, K.J.; Liu, K.; Hoogenboom, G.; White, J.W.; Smith, W.; Jegu, G.; Grant, B.; Crépeau, M.; Shang, J.; et al. Simulating the development and growth of lentil using the CSM-CROPGRO model. *Agron. J.* **2024**, *116*, 2391–2410. [\[CrossRef\]](#)
8. Ferreira, T.B.; Shelia, V.; Porter, C.; Cadena, P.M.; Cortasa, M.S.; Khan, M.S.; Pavan, W.; Hoogenboom, G. Enhancing crop model parameter estimation across computing environments: Utilizing the GLUE method and parallel computing for determining genetic coefficients. *Comput. Electron. Agric.* **2024**, *226*, 109349.
9. Guo, E.; Wang, Y.; Jirigala, B.; Jin, E. Integration of remote sensing and field observations in evaluating DSSAT model for estimating maize and soybean growth and yield in Maryland, USA. *Agronomy* **2023**, *13*, 1540. [\[CrossRef\]](#)
10. 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.; et al. *Decision Support System for Agrotechnology Transfer (DSSAT)*, Version 4.8.5; DSSAT Foundation: Gainesville, FL, USA, 2024.
11. Çığ, F.; Erman, M.; El Sabagh, A.; Ceritoglu, M.; Soysal, S.; Uçar, Ö.; Ceritoglu, F.; Kılıç, R.; Özek, R. Screening of lentil genotypes during germination and early growth stages under PEG-induced drought stress. *Chil. J. Agric. Res.* **2024**, *84*, 326–337. [\[CrossRef\]](#)
12. Ceritoglu, M.; Erman, M. Effect of vermicompost application at different sowing dates on some phenological, agronomic and yield traits in lentil. *J. Int. Environ. Appl. Sci.* **2020**, *15*, 158–166.
13. 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.; et al. The DSSAT crop modeling ecosystem. In *Advances in Crop Modelling for a Sustainable Agriculture*; Boote, K.J., Ed.; Burleigh Dodds Science Publishing: Cambridge, UK, 2019; pp. 173–216.
14. Dlugokencky, E.; Tans, P. Trends in Atmospheric Carbon Dioxide. 2020. Available online: <https://gml.noaa.gov/ccgg/trends/> (accessed on 25 May 2026).
15. Li, Z.; Jin, X.; Yang, G.; Xu, X.; Wang, J.; Lan, Y. Effects of different observed datasets on the calibration of crop model parameters with GLUE: A case study using the CROPGRO-Soybean phenological model. *PLoS ONE* **2024**, *19*, e0302098. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Willmott, C.J. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **1982**, *63*, 1309–1313. [\[CrossRef\]](#)
17. R Core Team. *R: A Language and Environment for Statistical Computing*, Version 4.5.1; R Foundation for Statistical Computing: Vienna, Austria, 2025. Available online: <https://www.R-project.org/> (accessed on 28 December 2025).
18. Wickham, H.; François, R.; Henry, L.; Müller, K.; Vaughan, D. dplyr: A Grammar of Data Manipulation, R Package Version 1.1.4. 2023. Available online: <https://CRAN.R-project.org/package=dplyr> (accessed on 1 April 2026).
19. Wickham, H.; Girlich, M. tidyr: Tidy Messy Data, R Package Version 1.3.1. 2024. Available online: <https://CRAN.R-project.org/package=tidyr> (accessed on 20 September 2025).
20. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
21. Slowikowski, K. ggrepel: Automatically Position Non-Overlapping Text Labels with ggplot2, R Package Version 0.9.5. 2024. Available online: <https://CRAN.R-project.org/package=ggrepel> (accessed on 18 December 2025).
22. Sehgal, A.; Sita, K.; Kumar, J.; Kumar, S.; Singh, S.; Siddique, K.H.M.; Nayyar, H. Effects of drought, heat and their interaction on the growth, yield and photosynthetic function of lentil (*Lens culinaris* Medikus) genotypes varying in heat and drought sensitivity. *Front. Plant Sci.* **2017**, *8*, 1776. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Sita, K.; Sehgal, A.; Bhandari, K.; Kumar, J.; Kumar, S.; Singh, S.; Siddique, K.H.; Nayyar, H. Impact of heat stress during seed filling on seed quality and seed yield in lentil (*Lens culinaris* Medikus) genotypes. *Front. Plant Sci.* **2017**, *8*, 1962. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Yao, N.; Wei, Y.; Jiang, K.; Liu, J.; Li, Y.; Ran, H.; Javed, T.; Feng, H.; Yu, Q.; He, J. Nonlinear water stress response functions can improve the performance of the DSSAT-CERES-Wheat model under water deficit conditions. *Agric. Water Manag.* **2025**, *307*, 109235. [\[CrossRef\]](#)
25. Lake, L.; Chenu, K.; Sadras, V.O. Accounting for soil water improves prediction of lentil phenology for improved frost and heat stress management. *Field Crops Res.* **2024**, *308*, 109287.
26. Vadez, V.; Soltani, A.; Sinclair, T.R. Crop simulation analysis of phenological adaptation of chickpea to different latitudes of India. *Field Crops Res.* **2013**, *146*, 1–9. [\[CrossRef\]](#)
27. Farooq, M.; Wahid, A.; Kobayashi, N.; Fujita, D.; Basra, S.M.A. Plant drought stress: Effects, mechanisms and management. *Agron. Sustain. Dev.* **2009**, *29*, 185–212. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.