Food Research International 233 (2026) 119004

Contents lists available at ScienceDirect

Food Research International

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

Soybean grain production and nutritional quality responses under elevated CO_2, high temperature, and drought

Janaina da Silva Fortirer [a,1], Adriana Grandis [a,1], Carmen Eusebia Palacios Jara [a], Débora Pagliuso [a], Leandro Francisco de Oliveira [d,e], Eveline Queiroz de Pinho Tavares [a], Lauana Pereira de Oliveira [a], Plínio B. Camargo [c], Eny Iochevet Segal Floh [d], Cibele M. Russo [b], Marcos S. Buckeridge [a,*]

[a] *Laboratory of Plant Physiological Ecology (LAFIECO), Department of Botany, Institute of Biosciences, University of São Paulo, São Paulo, SP, Brazil*
[b] *Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, SP, Brazil*
[c] *Laboratory of Isotopic Ecology, Centro de Energia Nuclear na Agricultura, University of São Paulo, Piracicaba, SP, Brazil*
[d] *Laboratory of Plant Cell Biology, Department of Botany, Institute of Biosciences, University of São Paulo, Rua do Matão, 277 São Paulo, SP, Brazil*
[e] *LigninLab, Department of Botany, Institute of Biosciences, University of São Paulo, Rua do Matão, 277 São Paulo, SP, Brazil*

ARTICLE INFO

Keywords::
Carbohydrates
Lipids
Nutritional Quality
Predictive Modeling
Proteins
Elevated CO2
Temperature
Soybean
Drought
Climate change

ABSTRACT

Soybean (*Glycine max* L. Merr.) is an important global crop that supplies protein and oil for the food and feed industries. However, its yield and nutritional quality are increasingly affected by climate change. This study combines experimental data with predictive modeling to assess how elevated CO_2 (eCO_2), high temperature, and drought (either alone or combined) influence soybean grain production and composition. Plants were grown under controlled conditions simulating future climate scenarios, and biochemical traits, including carbohydrates, proteins, lipids, and amino acids, were analyzed. Generalized linear models (GLMs) and machine learning algorithms (XGBoost, CatBoost) were used to predict yield and quality responses based on early–stage biomass data. Elevated CO_2 increased grain production up to 142%, while high temperature and drought reduced yield by 91% and 60%, respectively. The combined "Triple Effect" (CO_2 + high temperature + drought) was evaluated through predictive modeling from experimentally validated dual-stress datasets (eCO_2 + temperature and eCO_2 + drought), as this specific three-factor combination was not validated experimentally. Model projections a potential 50% increase in grain production, 35% soluble sugars, and a 175% rise in amino acid content, accompanied by reductions in starch (−20%) and protein (−6%). Elevated CO_2 may partially offset stress, while inducing metabolic shifts that could increase productivity, but alter grain nutritional quality. The integrated experimental with modeling framework highlights the importance of early physiological indicators and predictive tools to anticipating yield–quality trade–offs and support development of soybean cultivars resilient to multifactorial climate stress under future environmental conditions.

Abbreviations: aCO_2, Ambient carbon dioxide; ANOVA, Analysis of variance; C/N, Carbon to nitrogen ratio; CatBoost, Categorical Boosting; CO_2, Carbon dioxide; DAE, Days after emergence; Drought, Water deficit treatment; eCO_2, Elevated carbon dioxide; eCO_2+Drought, Elevated CO_2 combined with drought; eCO_2+Temp, Elevated CO_2 combined with high temperature; FACE, Free-Air CO_2 Enrichment; FAME, Fatty acid methyl ester; FAO, Food and Agriculture Organization; GC-FID, Gas chromatography with flame ionization detector; GLMs, Generalized Linear Models; HPAEC-PAD, High-performance anion-exchange chromatography with pulsed amperometric detection; OTC, Open-top chamber; WUE, Water-use efficiency; RICS, Remote Integrated Control System; RMSE, Root mean square error; SPAR, Soil-Plant-Atmosphere Research; Temp, High temperature treatment; Triple Effect, Combined elevated CO_2, high temperature, and drought; USDA, United States Department of Agriculture; XGBoost, Extreme Gradient Boosting..

* Corresponding author.
E-mail address: msbuck@usp.br (M.S. Buckeridge).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.foodres.2026.119004
Received 4 January 2026; Received in revised form 15 March 2026; Accepted 16 March 2026
Available online 18 March 2026

1. Introduction

Soybeans (*Glycine max* L. Merr.) are important to global food systems, significantly contributing to nutrition, bioenergy, and biotechnological innovations (Palacios et al., 2019; Wu et al., 2018). It is the world's leading grain crop, providing about 29% of the world's vegetable oil and serving as a major source of protein (SoyStats, 2023). Brazil is the largest producer (157 Mt – million tons), accounting for 39% of global production, followed by the United States (113 Mt – 28%), Argentina (50 Mt – 13%), and China (20 Mt – 5%) (USDA, 2024). In 2021, Brazil's soybean productivity (3.5 ton/ha) exceeded the global average (2.8 ton/ha), its position as the top exporter with 86 million tons (ANEC, 2022; CONAB, 2021; USDA, 2024). However, soybean production faces increasing threats under climate change scenarios, particularly elevated CO_2 (eCO_2), high temperatures, and drought, all of which can disrupt both production and quality. This, in turn, threatens global food security and economic stability (Beck, 2013).

Climate change introduces multiple interconnected stressors that impact soybean growth, productivity, and grain composition. Elevated atmospheric CO_2 generally boosts photosynthesis and biomass accumulation by improving carbon fixation and water–use efficiency (Digrado et al., 2024; Keenan et al., 2023; Thenveettil et al., 2024). Several studies have demonstrated that soybeans grown under eCO_2 increased net photosynthesis, resulting in higher vegetative biomass and pod production (Gong et al., 2025; Zheng et al., 2020). However, these physiological improvements are often paired with reductions in grain size, seed weight, and nutritional quality. Lower levels of protein, minerals, and starch have been observed under eCO_2, due to carbon–nitrogen dilution effects and changes in source–sink dynamics (Ekele et al., 2025; Machado et al., 2025; Poudel et al., 2023; Zipper et al., 2016). Such shifts in composition, balancing productivity with grain quality, raise concerns about the nutritional sufficiency of soy–based foods in a high–CO_2 future.

Temperature and water availability are major climate–related stressors affecting soybean yield and seed quality (Palacios et al., 2019). Soybeans grow best between 20 °C and 25 °C, but prolonged exposure to temperatures above 29 °C decreases pod numbers and seed weight. Temperatures over 35 °C severely disrupt flowering, pod set, and grain filling (Cai et al., 2020; Mourtzinis et al., 2015; Onat et al., 2017). Recent studies show that each 1 °C rise in average temperature can reduce yield by about 17%, depending on cultivar and growth stage (Thenveettil et al., 2024; Yang et al., 2023). High–temperature stress also changes the seed's biochemical makeup by increasing lipid saturation and decreasing protein and starch levels, which harms both technological and nutritional quality (Gong et al., 2025; Machado et al., 2025).

Water deficit is another major factor limiting soybean productivity and composition. Drought stress decreases photosynthetic efficiency, nitrogen fixation, and assimilate transport, resulting in smaller grains and yield losses of up to 40% (Gray et al., 2016; Poudel et al., 2023). Severe drought during reproductive stages disrupts carbohydrate and amino acid metabolism, reducing seed protein and oil levels (Ekele et al., 2025). Therefore, soybean yield and quality are highly vulnerable to changes in rainfall and soil moisture throughout the growing season (Sentelhas et al., 2015). Globally, climate–induced stressors have caused an estimated 21% decline in agricultural productivity since 1961, with tropical regions like Latin America being disproportionately impacted (Ortiz-Bobea et al., 2021).

Most studies assessing plant responses to environmental stressors have traditionally focused on individual factors or dual interactions, such as CO_2 x temperature or eCO_2 x drought (Ainsworth & Lemonnier, 2018; Ainsworth & Long, 2021; Fortirer et al., 2023; Fu et al., 2015; Pascual et al., 2022; Yoldi-Achalandabaso et al., 2025). Experimental evidence indicates that combined heat and drought stress can cause substantially greater damage to crops than individual stressors, due to compounded limitations on photosynthesis, metabolism, and reproductive development (Jiang et al., 2025; Pascual et al., 2022; Peláez-Vico et al., 2024; Zandalinas et al., 2024). Although elevated CO_2 may partially offset stress effects, its mitigation potential depends on species and environmental context (Shanker et al., 2022; Wang & Liu, 2021). Combining controlled experiments with predictive modeling has been recognized as a strategy for improving climate impact assessments, as empirical data provides mechanistic understanding while models enable extrapolation to future climate scenarios (Wang et al., 2017; Ewert et al., 2007). However, relatively few studies have investigated the simultaneous interaction of elevated CO_2, increased temperature, and drought within a unified experimental modeling framework, particularly in soybean (Yoldi-Achalandabaso et al., 2025).

Biomass accumulation serves as an early indicator of yield potential and changes in grain composition, offering valuable insights for predicting soybean performance under future climate scenarios in regions as Brazil (Bui et al., 2024; Monteiro et al., 2022). Recent progress in spatialized and process–based crop modeling has allowed researchers to expand these plant physiological responses from experiments to regional and national scales, supporting precision agriculture and sustainable food production (Pasquel et al., 2022). Climate projections forecast increases in both temperature extremes and drought frequency across major soybean–producing regions (Assad & Assad, 2024; FAO, 2023; Marengo et al., 2017). These conditions intensify soybean sensitivity to water deficits, leading not only to lower yields but also to changes in grain composition, especially in protein, lipid, and carbohydrate fractions (Ekele et al., 2025; Gray et al., 2016). Combining experimental data with predictive models that account for these combined stress responses is therefore essential for guiding breeding programs and maintaining nutritional quality in global food systems.

The complex interactions among climate variables, such as CO_2 levels, temperature, and water availability, call for robust modeling to forecast their combined effects on crop productivity and grain composition. Generalized Linear Models (GLMs) extend traditional regression by handling non–normal response distributions and complex interactions among predictors (Nelder & Wedderburn, 1972). They have been effectively used in agricultural and food research to estimate physiological traits, such as seed vigor, germination time, and uniformity, which are directly connected to seed quality and post–harvest performance (Jardim Amorim et al., 2021). Their flexibility and strong statistical foundation make them well–suited for modeling biological data that varies and depends on multiple factors. Recent research shows that GLMs can outperform simpler empirical models in predicting yield and quality outcomes, achieving lower root mean square error (RMSE) values and greater stability in predictions under different environmental conditions (Shastry et al., 2017).

Predictive models that incorporate key environmental stressors, such as eCO_2, high temperature, and drought, are important tools for predicting climate–driven effects on crop productivity and grain composition. These models can assist decision–making in agriculture and food security by identifying thresholds where yield and quality decline (Jin et al., 2017). However, most current approaches assess stressors separately, ignoring their complex, nonlinear interactions (Jin et al., 2017; Zandalinas & Mittler, 2022). Machine learning methods provide new opportunities to address these limitations by capturing complex relationships between environmental variables and plant physiological responses. Among the most reliable ensemble techniques, Extreme Gradient Boosting (XGBoost) and Categorical Boosting (CatBoost) have demonstrated superior predictive accuracy in modeling structured, multivariate datasets. XGBoost is well–known for efficiently modeling nonlinear dependencies and interactions (Chen & Guestrin, 2016; Li et al., 2023), while CatBoost enhances model generalization through ordered boosting and optimized handling of categorical features (Hancock & Khoshgoftaar, 2020; Prokhorenkova et al., 2018).

This study aimed to: (i) assess the predictive accuracy GLMs using early–stage physiological and morphological traits measured at 60 days after emergence (DAE) to forecast soybean grain yield and nutritional quality at harvest (125 DAE) under different environmental conditions;

and (ii) model the combined effects of eCO_2, high temperature, and drought—referred to as the "Triple Effect"—on soybean grain production and composition by integrating experimental datasets from dual–-stress treatments (e.g., eCO_2 + temperature or eCO_2 + drought). To our knowledge, this is the first report to present an integrated approach that connects experimental plant physiology with predictive analytics (via GLM and machine learning), providing a new framework for understanding how multifactorial climate stresses (Triple Effect) influence both productivity and grain nutritional quality in soybeans.

2. Materials and methods

2.1. Plant material and experimental design

Soybean (*Glycine max* L. Merr.) plants of the cultivar 'MG/BR–46 Conquista' (Embrapa, Brasília, Brazil) were grown under controlled conditions using open–top chambers (OTCs) in São Paulo, Brazil (23°34′1"S, 46°43′49″W) from September 2018 to February 2019. Seeds were germinated at 25 °C under a 12 h photoperiod and transferred to 5L pots containing a standardized substrate. Fifteen days after germination, seedlings were transferred to OTCs and exposed to ambient or elevated CO_2 concentrations, marking the start of the experiment (time zero, 0 DAE – days after emergence). Elevated CO_2 levels were set at 800 ppm (parts per million), which is twice the measured ambient concentration.

Two independent experiments were conducted in the same period. The first evaluated the combined effects of elevated CO_2 and increased temperature across four treatments: ambient CO_2 (aCO_2, 400 ppm), elevated CO_2 (eCO_2, 800 ppm), elevated temperature (Temp) (+5 °C above ambient), and their combination (eCO_2+Temp). The temperature increase (+5 °C) was applied relative to the ambient temperature conditions and was maintained continuously throughout the 24–h cycle. The eCO_2 concentration of 800 ppm and magnitude of +5 °C was selected to simulate projected warming scenarios under high–emission pathways (e.g., SSP5–8.5), which predict regional temperature increase of approximately 3–6 °C by the end of the 21st century in major growing regions (Ainsworth & Long, 2021; Zhao et al., 2017; IPCC, 2021; Ruiz-Vera et al., 2013). Such temperature increases are consistent with projections for subtropical and tropical agricultural regions, where warming is expected to impact crop physiology and productivity (IPCC, 2021). The second experiment evaluated the effects of water deficit and elevated CO_2 using four treatments: ambient CO_2 (aCO_2, 400 ppm), elevated CO_2 (eCO_2, 800 ppm), drought (Drought), and their combination (eCO_2+Drought). Drought treatment was imposed on plants after flowering (45 DAE) under both aCO_2 and eCO_2 conditions by suspending irrigation, followed by restricted watering (100 mL every 3 days) until physiological maturity (125 DAE), simulating progressive soil water conditions commonly observed under climate change scenarios (IPCC, 2021; Lesk et al., 2016)

All plants received 100 mL of Hoagland nutrient solution weekly (Hoagland et al., 1993). Microclimatic variables such as temperature, humidity, and CO_2 concentration were continuously monitored using automated sensors and recorded in the Remote Integrated Control System (RICS) software.

Ten plants per treatment were randomly harvested destructively at two different times: 60 and 125 days after emergence (DAE). Leaves, stems, roots, and pods were collected separately, freeze–dried, and grains were ground for biochemical analysis. The dried material from each plant organ was used to calculate biomass and grain production. The grains underwent biochemical evaluation for quality traits, including soluble sugars, starch, amino acids, lipids, carbon, and nitrogen content.

2.2. Non–structural carbohydrates in grain

Soluble sugars were measured from 10 mg of pulverized soybean grains using four sequential extractions with 80% (v/v) ethanol at 80 °C, as described by Arenque et al. (2014). All supernatants were centrifuged at 14,000 rpm at 4 °C and vacuum–dried (Savant SC 250 EXP, Thermo Scientific, Asheville, NC, USA). The residues were resuspended in ultrapure water, and pigments were removed using chloroform. All analyses involved 10 individuals per treatment. The aqueous fraction containing soluble sugars was analyzed by high–performance anion–exchange chromatography with pulsed amperometric detection (HPAEC-PAD; ICS 3000 system, Dionex–Thermo Fisher Scientific, Sunnyvale, CA, USA) equipped with a CarboPac PA1 column. Elution was performed isocratically for 27 min with 150 μM sodium hydroxide, and sugars (glucose, fructose, sucrose, and raffinose) were identified and quantified using external standards.

For starch determination, the dried pellet from ethanol extraction was sequentially digested with α–amylase (120 U/mL, Megazyme) in 10 mM MOPS buffer (pH 6.5) at 75 °C for 1 h, followed by amyloglucosidase (30 U/mL) in 100 mM sodium acetate buffer (pH 4.5) at 50 °C for 1 h. Glucose released from starch hydrolysis was measured using a coupled enzymatic assay containing glucose oxidase (1,100 U/mL), peroxidase (700 U/mL), 4–aminoantipyrine (290 μmol/L), and phenol (50 mM, pH 7.5). Absorbance was read at 490 nm after incubation at 30 °C for 15 min, and starch content was calculated from a glucose standard curve (Sigma) following Amaral et al. (2007) and Arenque et al. (2014).

2.3. Lipid extraction and fatty acid composition

Total lipids were extracted from 50 mg of dry soybean grain through three consecutive extractions with 1 mL of hexane at 50 °C for 30 min under constant agitation. The combined hexane extracts were evaporated to dryness, and lipid content was measured gravimetrically. The recovered oil was resuspended in Milli–Q water and transesterified to produce fatty acid methyl esters (FAMEs) following the method of Christie and Hutton (1993).

For transesterification, fatty acids were treated with 5% methanolic H_2SO_4 and toluene at 80 °C for 1 h. The reaction mixture was then homogenized with 0.5 M NaCl and dichloromethane, and the organic (lower) phase was separated, washed sequentially with dichloromethane and 0.05 M NaCl, and dried over anhydrous Na_2SO_4. The purified FAMEs were resuspended in hexane for chromatographic analysis.

Fatty acid profiles were analyzed using gas chromatography with a flame ionization detector (GC–FID; model 6850, Agilent Technologies). The method employed an HP–Innowax fused silica capillary column (cross–linked polyethylene glycol, 30 m × 320 μm × 0.50 μm film thickness). The oven temperature program was: start at 150 °C for 1 min, then increase at 15 °C min^{-1} to 225 °C, followed by a 5 °C min^{-1} to 260 °C, and hold at 270 °C for 7 min. Fatty acids were identified and quantified by comparing retention times and peak areas with those of analytical FAME standards (Sigma–Aldrich), using 10 individuals per treatment.

2.4. Carbon, nitrogen, C/N ratio, and total protein in grain

Elemental carbon and nitrogen contents were measured from 1.5 mg of finely ground soybean grain in 10 samples per treatment. Samples were placed in tin capsules and analyzed by combustion using an elemental analyzer (Carlo Erba 1110, Rodano, Italy) connected to an isotope ratio mass spectrometer (Finnigan Delta Plus, Thermo Fisher Scientific). The resulting CO_2 and N_2 gases were quantified to determine the total carbon and nitrogen levels, expressed as percentages of dry weight. Stable isotope ratios ($\delta^{13}C$ and $\delta^{15}N$) were calculated relative to international standards using the following equation:

$$13C \text{ or } 15N = \frac{(R\ sample - R\ standard)}{R\ standard} \text{ x } 1000 \quad (1)$$

Where *R* denotes the isotopic ratio of $^{13}C/^{12}C$ or $^{15}N/^{14}N$ for the sample and the reference standard. The C/N ratio was obtained directly from carbon and nitrogen concentrations. Total protein content was estimated from total nitrogen using a conversion factor of 6.25, as recommended for soybean (Mariotti et al., 2008).

2.5. Experimental data analysis

Data from the two independent experiments, temperature × CO_2 and drought × CO_2, were analyzed separately, considering 10 individuals per treatment and harvesting period (60 and 125 DAE). The data were examined for linearity, normality, and homoscedasticity to verify the model assumptions. Treatments conducted during the same sampling period were compared using one–way analysis of variance (ANOVA), followed by Tukey's post hoc test for mean separation ($P<0.05$), to evaluate the growth parameters and biochemical variables related to plant and grain production and quality. This approach allowed the assessment of main effects and their interactions under different environmental conditions. The analyses were performed using R software version 4.4.3 (R Core Team, 2024).

2.6. Experimental data used in models and data pre–processing

The total dry biomass measured 60 days after the start of the experiment (DAE) (Table S1) was used as the primary input variable for model development. The responses observed at 125 DAE, including grain yield, non–structural carbohydrates, lipids, proteins, and carbon/nitrogen ratios, were used for model validation. Experimental scenarios were designed to assess the isolated and combined effects of aCO_2, eCO_2, Temp, and Drought.

Categorical variables were converted using one–hot encoding, generating binary dummy variables for each treatment (aCO_2, eCO_2, Temp, eCO_2+Temp, Drought, and eCO_2+Drought). To maintain data balance and comparability among treatments, random subsampling was applied to the aCO_2 and eCO_2 groups, bringing all treatment groups to $n = 10$ replicates. Before modeling, the data were checked for linearity, normality, and homoscedasticity to ensure they met model assumptions. Spearman's rank correlation was used to assess relationships among predictors and response variables (De Winter et al., 2016).

For training machine learning models, all datasets were curated, cleaned, and standardized. Continuous features were normalized using the min–max scaling technique to transform each variable into the [0, 1] range according to Equation (2):

$$V' = v - v_{min}/v_{max} - v_{min} \quad (2)$$

Where v represents the original feature value, V' the normalized value, and v_{min} and v_{max} the minimum and maximum observed values, respectively. This normalization ensured comparable feature scales and improved model convergence.

Categorical variables were encoded with one–hot encoding, and the entire dataset was randomly split into training (80%) and testing (20%) subsets, keeping a stratified representation of all treatments to prevent sampling bias.

2.7. Generalized linear models (GLMs)

GLMs were used to analyze the relationships between environmental treatments and the observed responses in soybean grain production and quality. Unlike traditional parametric models that assume normally distributed errors, GLMs allow response variables to follow different probability distributions, such as the Gamma distribution, offering greater flexibility for modeling complex biological data (Nelder & Wedderburn, 1972). The GLM was formulated as shown in Equation (3).

$$g(E(y)) = \beta_0 + \beta_1\,(eCO_2) + \beta_2\,(Temp) + \beta_3\,(Drought) + \beta_4\,(eCO_2 + Temp) + \beta_5\,(eCO_2 + Drought) \quad (3)$$

Where y represents the expected value of the dependent variable, g is the link function that relates the mean response to the linear predictors, β_0 is the intercept, and β_1–β_5 are the coefficients associated with each treatment effect.

Model predictions were used to estimate the relative percentage change in production and quality parameters under each treatment compared to the aCO_2 baseline. Model performance was assessed by comparing predicted and observed values using the root mean square error (RMSE), as described by Loague and Green (1991) (Equation 4):

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad (4)$$

where n is the number of treatments, $\hat{y}_i$ represents the predicted value, and y_i the observed value.

To further investigate the direct and indirect effects of the variables eCO_2, temperature, and drought on grain yield and nutritional quality, a path analysis framework based on GLMs was used, following Streiner (2005). All continuous variables were standardized using z–score normalization to allow comparability across different measurement scales (Di Paola et al., 2023), as shown in Equation (5).

$$Z = (x - \mu)/\sigma \quad (5)$$

where x is the original value, μ is the mean, and σ *is* the standard deviation. Data processing and modeling were conducted in R version 4.4.3 (R Core Team, 2024) using the *dplyr* package (Wickham et al., 2014).

2.8. Machine Learning modeling with XGBoost and CatBoost

Machine learning algorithms were used to develop predictive models for soybean grain yield and nutritional quality under various environmental stress conditions. Two ensemble gradient–boosting techniques were employed: XGBoost (Chen & Guestrin, 2016; Li et al., 2023) and CatBoost (Nishat et al., 2025; Prokhorenkova et al., 2018). These algorithms were chosen for their ability to manage nonlinear relationships and complex feature interactions while reducing overfitting through regularization. CatBoost uses an objective function similar to that of XGBoost (Equation 6):

$$\mathscr{L} = \sum_{i=1}^{n} l\,(yi, \hat{y}i) + \Omega(f) \quad (6)$$

where $l(y, \hat{y})$ denotes the loss function measuring the difference between observed (y) and predicted ($\hat{y}$) values at 125 DAE, and $\Omega(f)$ is a regularization term that controls the model's complexity.

Model optimization was carried out through systematic hyperparameter tuning. For XGBoost, the main parameters adjusted included learning rate, maximum tree depth, number of estimators, and minimum child weight. For CatBoost, the parameters optimized were number of iterations, tree depth, learning rate, and L2_leaf_reg (Table S2). Randomized grid search and five–fold cross–validation were used to ensure robustness during training, and early stopping criteria were applied to prevent overfitting (Kohavi, 1995; Prechelt, 1998).

Model performance was assessed using the RMSE (Loague & Green, 1991) between predicted and experimental values. The trained models were then employed to estimate soybean grain yield and compositional quality under isolated and combined climate scenarios (eCO_2, temperature, and drought). This predictive framework offers an integrative approach to forecast yield–quality trade–offs and aids decision–making for sustainable food production in future climate conditions.

3. Results

3.1. Microclimatic data

Throughout the experimental period, environmental conditions in the OTCs remained stable and within expected ranges for each treatment. In the aCO_2 and eCO_2 treatments, the average air temperature ranged from 25 to 26 °C, and the relative humidity ranged from 68% to 71% (Fig. S1). The high–temperature treatment maintained an average of 29 °C, a +5 °C increase over ambient conditions, with an average relative humidity of 57% (Fig. S1). In the drought experiment, mean air temperature was around 24 °C, with relative humidity ranging from 59% to 78% (Fig. S1). These controlled microclimatic conditions ensured reliable differentiation among treatments and consistent simulation of the intended environmental stressors.

3.2. Soybean grain production and quality under isolated and combined effects of elevated CO_2, temperature, and drought

Grain production consistently increased under eCO_2 compared to aCO_2 in both independent experiments (drought and temperature). Grain biomass reached 13.9 and 19.0 g per plant under eCO_2, exceeding the corresponding aCO_2 values (10.6 and 4.8 g) (Fig. 1a–b). Variations among ambient replicates were mainly due to differences in sunlight exposure, as plants in Drought conditions received higher irradiance than those in the temperature–controlled OTCs. Despite this, the trend of yield increase under eCO_2 remained consistent across both experiments.

High temperature significantly reduced grain formation, leading to a 91% decline in grain biomass (0.63 g) compared to aCO_2. Conversely, the combined eCO_2+Temp treatment alleviated temperature stress, resulting in a 59% increase in yield (16.9 g) relative to aCO_2 (Fig. 1a). High temperature markedly increased soluble sugar accumulation, with glucose and fructose concentrations rising about fourfold compared to aCO_2 (Figure 2a–b). Raffinose content also increased notably under eCO_2, from 1.74 $\mu g.mg^{-1}$ in aCO_2 to 7.57 $\mu g.mg^{-1}$ (+335%) (Fig. 2c), while sucrose levels remained stable (Fig. 2d). Elevated CO_2 and eCO_2+Temp promoted greater carbohydrate synthesis and allocation to developing grains (Fig. 2).

Under drought conditions, glucose and fructose concentrations increased by 52% (2.76 $\mu g.mg^{-1}$) and 76% (1.20 $\mu g.mg^{-1}$), respectively, compared with aCO_2 (Fig. 2f, g). However, raffinose peaked under drought (8.14 $\mu g.mg^{-1}$) and eCO_2+Drought (10.5 $\mu g.mg^{-1}$), while eCO_2 alone showed the lowest level (2.14 $\mu g.mg^{-1}$) (Fig. 2h). Sucrose content increased significantly (+39%) only under eCO_2+Drought (39.3 $\mu g.mg^{-1}$) (Fig. 2i), whereas starch concentration was higher under eCO_2 (24.2 $\mu g.mg^{-1}$) compared to aCO_2 (16.5 $\mu g.mg^{-1}$) (Fig. 2j).

The total lipid content decreased by 8% at Temp compared to the aCO_2 and eCO_2 treatments (Fig. 3a). However, in the Drought, oil content did not decline in comparison with aCO_2 with an increase of 13% in eCO_2 (Fig. 3b), possibly due to acceleration in grain store development. The fatty acid composition varied significantly, reflecting this developmental mismatch.

Elevated temperature treatments promote an increase in saturated (palmitic and stearic) and monounsaturated (oleic) fatty acids, and the eCO_2+Temp condition maintained this increase (Fig. 3c). In particular, levels of palmitic, stearic, and oleic acids rose by 14%, 7%, and 28%, respectively, under high–temperature conditions, reflecting shifts in lipid metabolism during grain development. The Temp and eCO_2+Temp treatments decreased the levels of polyunsaturated fatty acids (linoleic and linolenic).

On the other hand, Drought treatment reduced the relative abundance of saturated (palmitic) and polyunsaturated (linoleic and linolenic) fatty acids. Still, oleic acid (monounsaturated) increased with Drought and eCO_2 (Fig. 3d). Higher proportions of unsaturated fatty acids were observed under aCO_2 and eCO_2. In contrast, Temp and Drought increased the relative abundance of saturated chains (Fig. 3c, d). In any case, both eCO_2+Temp and eCO_2+Drought reduced linoleic acid, the major fatty acid in soybean oil, indicating a preference for the accumulation of the monounsaturated fatty acid, oleic acid.

Temperature and drought, both alone and combined with eCO_2, increased amino acid accumulation (Fig. 4a, b). The eCO_2+Temp treatments produced the highest amino acid levels, five times greater than aCO_2, followed by Temp (four times), Drought (three times), and eCO_2+Drought (two times). Despite these differences, total nitrogen and crude protein levels showed no significant variation among treatments (Figs. 4C–D, 5A).

Carbon concentration reached its maximum under eCO_2, averaging 48–50% (Figs. 5B, 5E). In contrast, high–temperature treatments showed a decline in both total carbon and the C/N ratio relative to their ambient counterparts (Fig. 5c). Elevated CO_2 enhances carbon assimilation and carbohydrate storage. At the same time, Tempand Drought treatments induce metabolic adjustments that modify sugar and amino acid profiles, with potential implications for grain nutritional quality.

3.3. Modeling soybean grain production and quality using early–stage biomass predictors (60 DAE) under climate stress conditions

To forecast soybean grain production and nutritional quality at 125 DAE and total plant biomass weights collected at 60 DAE were used as predictors for each experimental treatment. GLMs were fitted to evaluate how physiological traits under early growth conditions explain grain biochemical responses, such as production, soluble sugars, starch, lipids, proteins, and amino acids, across different climate change scenarios.

GLMs showed strong agreement between observed and predicted values, with RMSE, indicating high predictive reliability. Grain

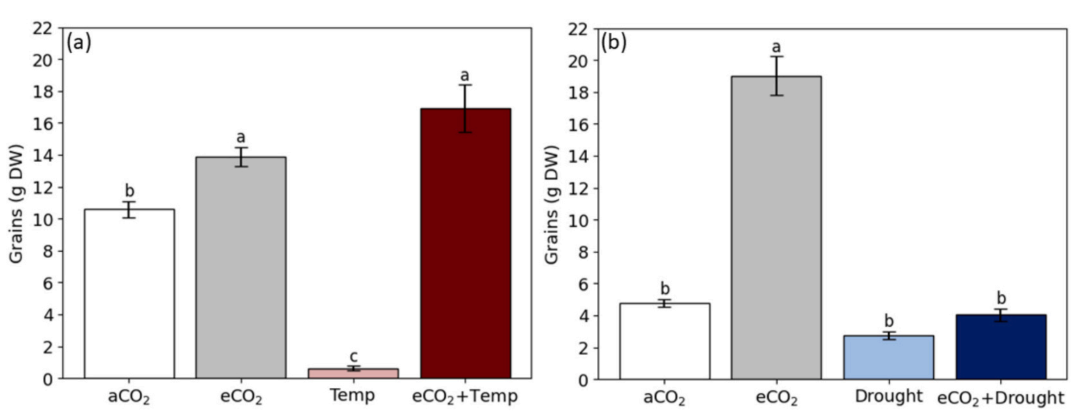
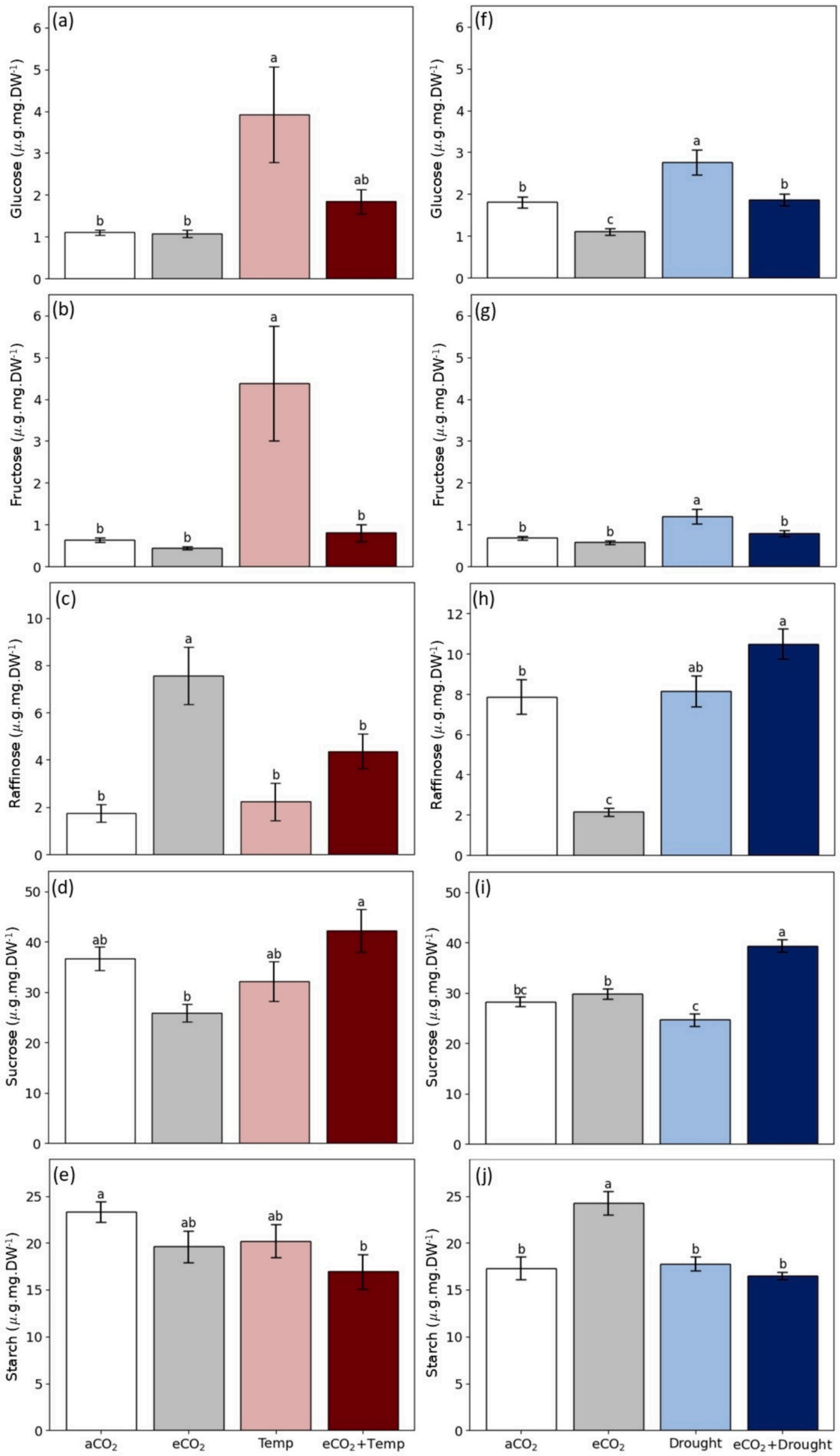


Fig. 1. Soybean grains dry weight (g) at 125 days after emergence under ambient, elevated CO_2, temperature, and drought conditions. Panel (a) shows treatments applied include aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C). Panel (b) including aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Drought (400 ppm CO_2 + watering reduction); and eCO_2+Drought (800 ppm CO_2 + watering reduction). Data are presented as means ± standard errors ($n = 10$). Different letters indicate significant differences among treatments according to Tukey's test ($p < 0.05$).



(caption on next page)

Fig. 2. Non-structural carbohydrates in soybean grains (μg.mg DW^{-1}) at 125 days under ambient, elevated CO_2, temperature, and drought conditions. The panels on the left (a–e) show temperature treatments, while those on the right (f–j) display drought treatments. Glucose (a, f), fructose (b, g), raffinose (c, h), sucrose (d, i), and starch (e, j) in mature grains harvested from soybeans. Treatments include [aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Drought (400 ppm CO_2 + watering reduction); and eCO_2+Drought (800 ppm + watering reduction)]. Data are shown as means ± standard errors (n=10). Different letters indicate significant differences among treatments, according to Tukey's test (p < 0.05).

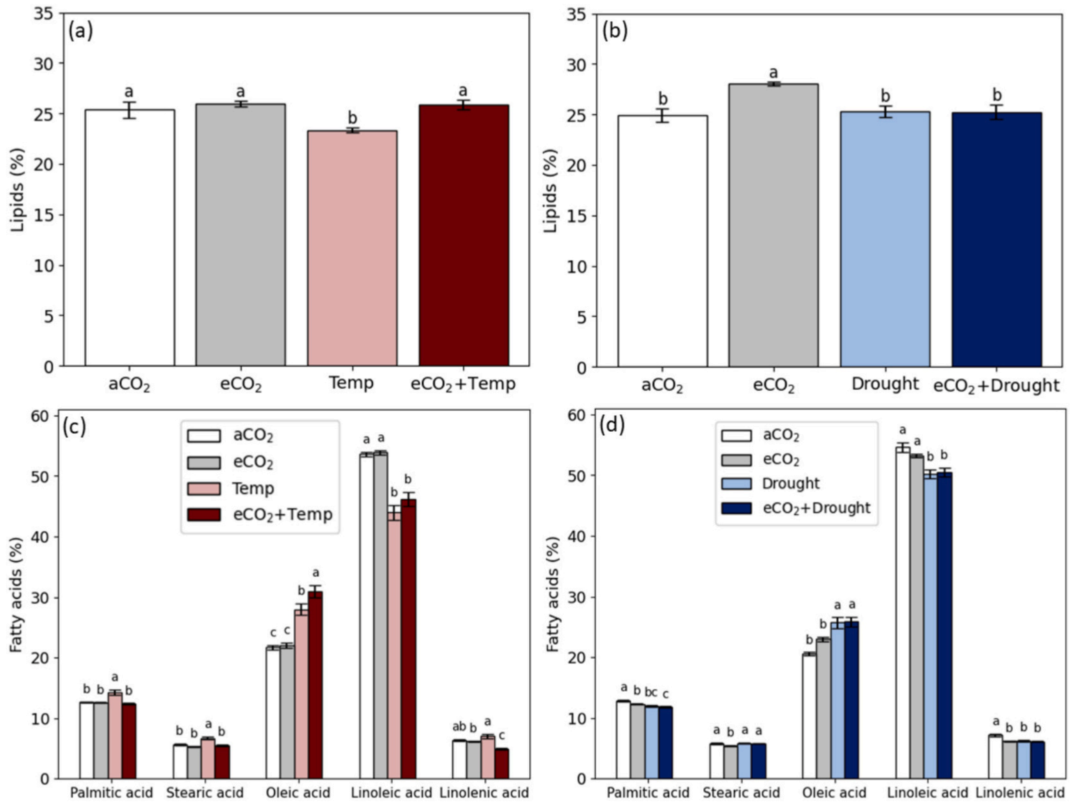


Fig. 3. Total lipids and fatty acids content (%) in soybean grains at 125 days after the experiment under ambient, elevated CO_2, temperature, and drought conditions. Panels on the left (a and c) show temperature treatments, while the panels on the right (b and d) show drought treatments. Fatty acids include palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid in mature grains harvested from soybeans. Treatments used include [aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Drought (400 ppm CO_2 + watering reduction); and eCO_2+Drought (800 ppm CO_2 + watering reduction)]. Data are presented as means ± standard errors (n = 10). Different letters indicate significant differences among treatments based on Tukey's test (p < 0.05).

production had the best fit under eCO_2 conditions, with observed and predicted values of 16.9 ± 1.01 g and 16.6 ± 0.79 g, respectively (RMSE = 3.85) (Table 1). Conversely, Drought caused significant reductions across all measured traits, confirming its strong negative impact on yield and grain biochemical composition (Table 1). Soluble sugars were significantly higher under eCO_2+Temp conditions (49.2 ± 3.8 μg·mg^{-1}), with model predictions closely matching experimental values (48.1 ± 2.7 μg·mg^{-1}; RMSE = 13.3). Amino acids showed the greatest sensitivity to environmental stressors, tripling under eCO_2+Temp (observed 290.4 ± 23.2 μmol·mg^{-1}; predicted 278.1 ± 15.5 μmol·mg^{-1}; RMSE = 85.6). In contrast, amino acid concentrations sharply decreased under Drought and eCO_2+Drought conditions (Table 1).

Path analysis diagrams (Fig. 6) were used to quantify and visualize the direct and indirect effects of environmental treatments on grain production and quality traits (Hallgren et al., 2019). Each arrow indicates a directional relationship, with standardized beta coefficients (β) showing the magnitude and direction of effects. Positive β-values represent stimulatory relationships, while negative coefficients imply inhibitory or antagonistic effects. Line thickness reflects statistical significance (p < 0.05), as detailed in Table 2.

The path analysis showed that total biomass at 60 DAE significantly predicted grain production at 125 DAE across all environmental conditions. Elevated CO_2 and eCO_2+Temp treatments had strong positive effects on grain production, while Temp, Drought, and eCO_2+Drought had negative effects (Fig. 6; Table 2). Total soluble sugars had a moderate negative relationship under Drought (β = −0.14), indicating a trade–off between sugar accumulation and productivity during water stress. Lipid content was positively associated only under eCO_2 (β = 0.08), suggesting increased carbon allocation toward lipid synthesis in high–CO_2 environments. Amino acid responses varied among treatments, with positive associations under eCO_2+Temp (β = 0.05) and negative relationships under eCO_2+Drought (β = −0.07) and eCO_2 (β = −0.06).

3.4. Machine learning prediction of soybean grain production and quality under the climate change "Triple Effect"

A comparative analysis of two ensemble learning algorithms, XGBoost and CatBoost, was conducted to predict soybean grain production and nutritional quality under combined elevated CO_2, high temperature, and drought conditions (Triple Effect). The predicted results were not experimentally validated. Model performance metrics are summarized in Table 3. XGBoost demonstrated superior predictive accuracy across all evaluated traits, with the lowest root mean square error

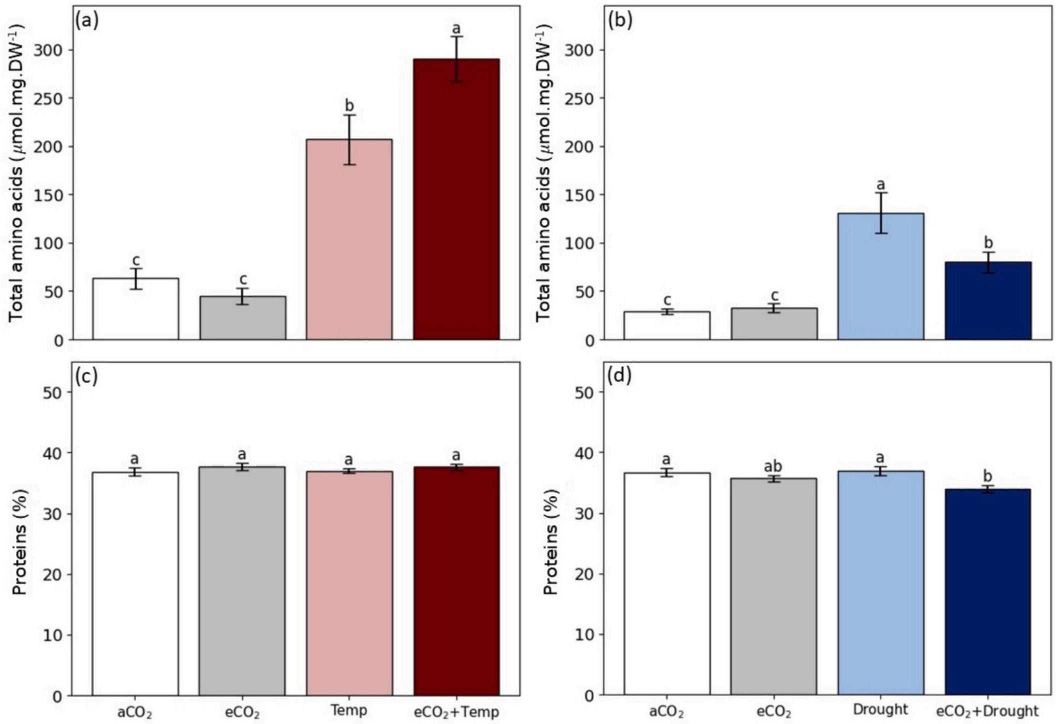


Fig. 4. Proteins and amino acids in soybean grains at 125 days after the experiment under ambient, elevated CO_2, temperature, and drought conditions. Total amino acids (a, b) and proteins (c, d) in mature grains harvested from soybeans. Treatments applied include [aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Drought (400 ppm CO_2 + watering reduction); and eCO_2+Drought (800 ppm CO_2 + watering reduction)]. Data are shown as means ± standard errors (n = 10). Different letters indicate significant differences among treatments based on Tukey's test ($p < 0.05$).

(RMSE = 0.04) for grain yield prediction, followed by CatBoost (RMSE = 1.50). For grain quality parameters, XGBoost also outperformed CatBoost, showing a greater ability to capture nonlinear and multivariate interactions among environmental stressors (Table 3).

Model–based projections showed that, compared to ambient conditions, the combined Triple Effect scenario resulted in a 50% increase in grain production, a 35% rise in soluble sugar concentration, and a remarkable 175% increase in total amino acid content. In contrast, starch and protein levels dropped by about 20% and 6%, respectively, while lipid concentrations stayed stable across treatments (Figure 7).

4. Discussion

Understanding the interactive effects of elevated CO_2, high temperature, and drought remains one of the most important challenges in predicting crop performance under future climate scenarios (Jin et al., 2017). By combining experimental data with predictive modeling, this study separated the effects of abiotic factors (whether isolated or combined) on soybean biomass, physiological traits, and grain biochemical composition. The mix of empirical measurements and modeling offers a strong framework for understanding complex plant responses and validating predictive approaches (Qiao et al., 2019; Wu et al., 2018).

Although many crop models simulate growth based on weather, soil, and management variables, they often lack physiological and biochemical constraints such as nutrient availability, carbon/nitrogen partitioning, and stress signaling pathways (Nendel et al., 2023). Our results show that eCO_2 enhances carbohydrate synthesis and carbon allocation, leading to increased biomass and yield. The observed increase in soluble sugars and amino acids indicates metabolic reprogramming under stress, aligning with previous reports that connect eCO_2 to osmotic regulation, reactive oxygen scavenging, and stress adaptation (Palacios et al., 2019; Zinta et al., 2018).

By comparing individual (eCO_2, Temp, Drought), dual (eCO_2 + Temp , or eCO_2 =+ Drought), and the combined predicted Triple Effect (eCO_2 with Temp, and Drought) treatments, our models showed that responses depend on the type and strength of environmental interactions. The GLMs approach effectively captured the linear and additive aspects of temperature and water stress, while the XGBoost model more accurately described nonlinear and hierarchical interactions among all three factors. This improved predictive ability supports a more comprehensive understanding of climate–driven impacts on grain production and nutritional quality (Figure 7). Because this study was conducted in OTCs, with a specific cultivar, the results may not fully represent plant responses under natural field environments. Therefore, further validation under field conditions is necessary to provide a more comprehensive understanding of plant performance under experimentaland agronomic conditions.

4.1. Anticipating grain production and quality under climate change

Modeling grain production and quality from early–stage experimental data offers a powerful way to predict how crops will respond to climate change. Several studies highlight the predictive value of early vegetative traits for final yield components, enhancing forecast accuracy while reducing the need for long–term field measurements (Van Eeuwijk et al., 2019). Our approach has proven effective, showing that early vegetative responses are strong indicators of final grain outcomes, especially under multifactorial climate stress, using GLMs modeling. Models show that early physiological performance reliably predicts final yield and compositional quality under complex climate stress factors. Elevated CO_2 partially lessens the negative impacts of temperature and drought on grain yield; however, it can also change biochemical partitioning, emphasizing the need to balance productivity and nutritional outcomes in future climate conditions.

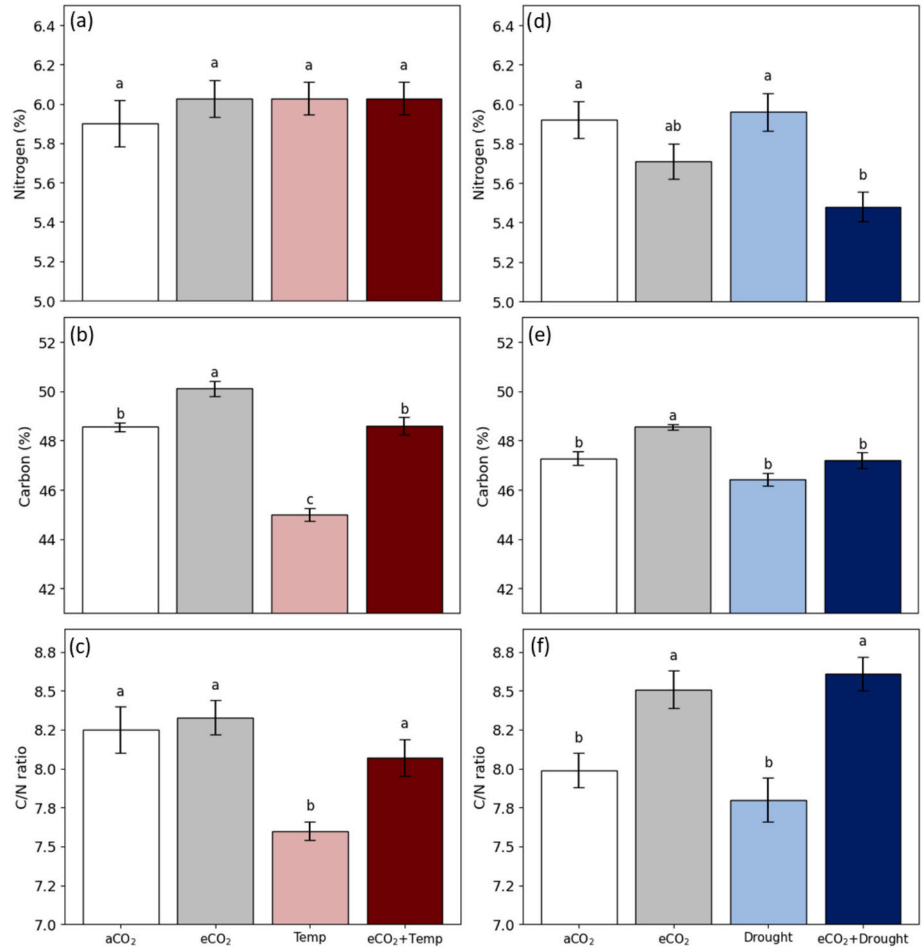


Fig. 5. Carbon and nitrogen content (%) in soybean grains at 125 days after the experiment under ambient, elevated CO_2, temperature, and drought conditions. Panels on the left (a-c) show temperature treatments, while panels on the right (d-f) display drought treatments. Carbon and nitrogen contents from soybean grains are represented by nitrogen (a, d), carbon (b, e), and carbon/nitrogen ratio (c, f) in mature grains harvested from soybeans. Treatments included: [aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); Drought (400 ppm CO_2 + reduced watering); eCO_2+Drought (800 ppm CO_2 + reduced watering)]. Data are shown as means ± standard errors (n = 10). Different letters indicate significant differences among treatments according to Tukey's test ($p < 0.05$).

Across all evaluated traits, including grain yield, soluble sugars, starch, lipids, proteins, and amino acids, the models showed strong agreement between observed and predicted values with RMSE (Table 1). Our findings align with previous research, which suggests that statistical models based on early physiological indicators can provide biologically meaningful estimates of crop productivity and nutritional quality under abiotic stress (Chenu et al., 2017). For example, under eCO_2 conditions, grain production reached 16.9 ± 1.01 g, while the GLMs predicted 16.6 ± 0.79 g (RMSE = 3.85), confirming the predictive ability of early biomass indicators, especially in carbon–enriched environments.

Amino acid concentrations increased nearly five–fold in eCO_2+Temp compared to aCO_2 conditions (Table 1). The model successfully reproduced this trend, capturing the metabolic adjustments linked to combined stress exposure. This finding aligns with results by Zinta et al. (2018), who reported that concurrent climate stressors influence amino acid metabolism as part of adaptive mechanisms to maintain osmotic balance and metabolic flexibility. We demonstrated the predictive ability of early–stage physiological and morphological data for estimating final grain yield and quality. This framework supports developing climate–resilient cultivars and enhances the capacity to predict potential yield losses (Banga & Kang, 2014; Khan et al., 2022). Furthermore, it highlights the usefulness of GLMs in integrating multiple environmental and physiological factors, thereby informing adaptive breeding and agronomic strategies under projected climate scenarios.

4.2. Validation of the model using out–of–data experiments

Model–based projections from GLMs revealed distinct impacts of eCO_2, Temp, Drought, and their interactions on soybean production and quality (Table 1). The model accurately reproduced the trend of yield increase under eCO_2, with predicted grain production (16.52 ± 0.96 g) closely matching the observed value (14.67 ± 0.39 g) (Table S3). Starch estimates were also consistent under aCO_2 (11.27 ± 0.87 μg mg^{-1} predicted vs. 11.54 ± 0.93 μg mg^{-1} observed) and eCO_2 (16.52 ± 0.96 μg mg^{-1} predicted vs. 14.67 ± 0.39 μg mg^{-1} observed). However, the model tended to overestimate production under combined eCO_2+Temp (20.58 ± 1.76 μg mg^{-1} predicted vs. 10.04 ± 1.26 μg mg^{-1} observed) and failed to capture the sharp decline in soluble sugars (15.68 ± 0.30 μg mg^{-1} predicted vs. 2.95 ± 0.54 μg mg^{-1} observed) (Table S3). These inconsistencies, especially the overestimation of protein content and the unexpected increase in starch biosynthesis under dual stress, suggest partial model overfitting and limitations related to simplified assumptions of source–sink balance and C/N partitioning. These issues likely arise from the limited dataset size, which restricts the representation of genotypic variability, developmental stages, and environmental heterogeneity.

Expanding the calibration dataset across different sites, seasons, and genetic backgrounds would improve the model's realism and broad applicability. Integrating data through meta–analysis of multiple sites

Table 1
Observed and predicted values of grain production and quality traits (soluble sugars, starch, lipids, proteins, and amino acids) under different treatments as estimated by generalized linear models (GLMs). Treatments include ambient CO_2 (aCO_2), elevated CO_2 (eCO_2), increased temperature (Temp), drought, and their combinations. Data is represent mean ± standard error of observed and predicted values, the root mean square error (RMSE), and the fitted model equation in the form y = α + βx, where α is the intercept and β the slope coefficient. Trait-specific intercepts (α) are provided above each variable (production, soluble sugars, starch, lipids, proteins, and amino acids).

Grain	Observed	Predicted	RMSE	Model (y = α+ βx)
Production (gDW)				*α = 1.67*
aCO_2	6.97 ± 0.78	6.62 ± 0.33	3.04	y = 0.638x + 0.172
eCO_2	16.9 ± 1.01	16.56 ± 0.79	3.85	y = 0.809x + 0.142
Temp	0.63 ± 0.15	0.65 ± 0.02	0.47	y = -0.116x + 0.118
Drought	2.78 ± 0.24	2.76 ± 0.09	0.74	y = -0.138x + 0.141
eCO_2+Temp	16.93 ± 1.48	16.24 ± 0.90	5.92	y = 0.879x + 0.154
eCO_2+Drought	4.04 ± 0.41	3.90 ± 0.39	1.39	y = 0.208x + 0.107
Soluble sugars (μg.mgDW^{-1})				*α = 40.28*
aCO_2	33.73 ± 1.78	32.32 ± 1.61	9.91	y = -0.838x + 0.569
eCO_2	34.26 ± 1.33	33.66 ± 1.61	5.76	y = -0.300x + 0.260
Temp	42.66 ± 2.23	42.32 ± 1.23	7.73	y = 0.268x + 0.599
Drought	28.59 ± 1.35	28.29 ± 0.98	5.01	y = -1.434x + 0.617
eCO_2+Temp	49.23 ± 3.77	48.05 ± 2.68	13.32	y = 0.495x + 0.297
eCO_2+Drought	42.00 ± 1.26	38.73 ± 0.21	11.41	y = 0.175x + 0.426
Starch (μg.mgDW^{-1})				*α = 18.96*
aCO_2	20.95 ± 1.52	20.25 ± 1.01	6.36	y = 0.201x + 0.400
eCO_2	22.18 ± 1.57	21.48 ± 1.03	6.59	y = 0.155x + 0.184
Temp	20.22 ± 1.78	19.92 ± 0.58	6.16	y = 0.126x + 0.399
Drought	17.77 ± 0.73	17.46 ± 0.60	3.55	y = -0.164x + 0.425
eCO_2+Temp	16.94 ± 1.86	16.87 ± 0.94	5.06	y = -0.83x + 0.185
eCO_2+Drought	16.51 ± 0.40	14.97 ± 1.56	5.25	y = -0.195x + 0.265
Lipids (%)				*α = 23.77*
aCO_2	25.51 ± 0.48	25.85 ± 1.24	4.63	y = 0.186x + 0.144
eCO_2	26.87 ± 0.35	26.33 ± 1.26	3.86	y = 0.159x + 0.066
Temp	23.34 ± 0.25	23.18 ± 0.67	1.87	y = -0.046x + 0.142
Drought	25.29 ± 0.58	25.01 ± 0.86	3.22	y = 0.186x + 0.158
eCO_2+Temp	25.88 ± 0.48	25.19 ± 1.40	4.41	y = 0.115x + 0.069
eCO_2+Drought	25.26 ± 0.68	23.34 ± 2.44	6.58	y = 0.145x + 0.103
Proteins (%)				*α = 34.31*
aCO_2	38.01 ± 0.61	37.19 ± 1.85	5.79	y = 0.421x + 0.171
eCO_2	37.00 ± 0.60	36.25 ± 1.74	5.50	y = 0.138x + 0.077
Temp	37.02 ± 0.39	36.70 ± 1.07	3.80	y = 0.304x + 0.170
Drought	36.91 ± 0.71	36.59 ± 1.26	3.63	y = 0.331x + 0.186
eCO_2+Temp	37.67 ± 0.51	36.56 ± 2.04	6.88	y = 0.179x + 0.082
eCO_2+Drought	33.95 ± 0.57	31.10 ± 3.25	9.63	y = -0.010x + 0.118
Amino acids (%)				*α = 115.85*
aCO_2	48.51 ± 8.56	44.35 ± 2.21	18.37	y = -8.17x + 2.403
eCO_2	44.84 ± 6.30	42.76 ± 2.05	17.17	y = -3.63x + 1.078
Temp	206.84 ± 25.82	204.04 ± 5.95	79.29	y = 10.356x + 3.970
Drought	130.72 ± 20.97	128.07 ± 4.44	63.23	y = 1.713x + 3.444
eCO_2+Temp	290.42 ± 23.23	278.09 ± 15.54	85.59	y = 9.826x + 2.462
eCO_2+Drought	79.97 ± 10.69	72.41 ± 7.58	33.56	y = -2.655x + 1.802

could refine the parameterization of key physiological processes. Additionally, combining the GLMs framework with transcriptomic or metabolomic datasets would help tighten predictions, reduce overfitting, and better capture metabolic feedback under combined stressors (Wang et al., 2024). Merging empirical measurements with modeling thus creates a strong platform for data validation and mechanistic understanding, as highlighted by Qiao et al. (2019), and Wu et al. (2019). Predictive modeling remains a useful tool for evaluating environmental effects on grain yield and quality, aiding the development of adaptive management strategies in response to projected climate change (Wang et al., 2018; Wang & Liu, 2021). The complex interactions among elevated CO_2, high temperature, and drought highlight the difficulty of accurately predicting crop responses under future conditions (Jin et al., 2017).

Future modeling efforts should explicitly include physiological and biochemical variables, such as nutrient availability, tissue susceptibility, and compositional quality, alongside traditional weather and management parameters (Nendel et al., 2023). Combining experimental evidence with predictive modeling not only improves accuracy but also clarifies how elevated CO_2 enhances carbon assimilation and nitrogen uptake, as well as how these benefits are limited under temperature and drought stress. The buildup of soluble sugars and amino acids indicates activation of stress–signaling and osmoprotective pathways that maintain metabolic flexibility under combined stresses (Heinemann & Hildebrandt, 2021; Palacios et al., 2019; Zinta et al., 2018).

4.3. Elevated CO_2 promotes grain yield gains but alters compositional and nutritional quality

Elevated concentration of atmospheric CO_2 enhances photosynthetic carbon assimilation, improves water-use efficiency (WUE), stimulates biomass accumulation, and grain yield in soybean (DaMatta et al., 2010; Digrado et al., 2024; Shanker et al., 2022). These physiological responses increase the availability of photoassimilates during seed development, promoting the accumulation of carbon storage compounds, such as lipids and carbohydrates (Leakey et al., 2009; Taub et al., 2008; Li et al., 2023). eCO_2 increased oil concentration, soluble sugars, and starch content (Fig. 3–7).

While carbon compounds increase under elevated CO_2, responses of protein and nitrogen metabolism are more variable. Legumes such as soybeans may partially offset nitrogen dilution effects through symbiotic nitrogen fixation, helping maintain nitrogen balance under elevated CO_2 (Myers et al., 2014). Reports show modest protein increases (1.5%) under elevated CO_2 (Digrado et al., 2024), while others observe no

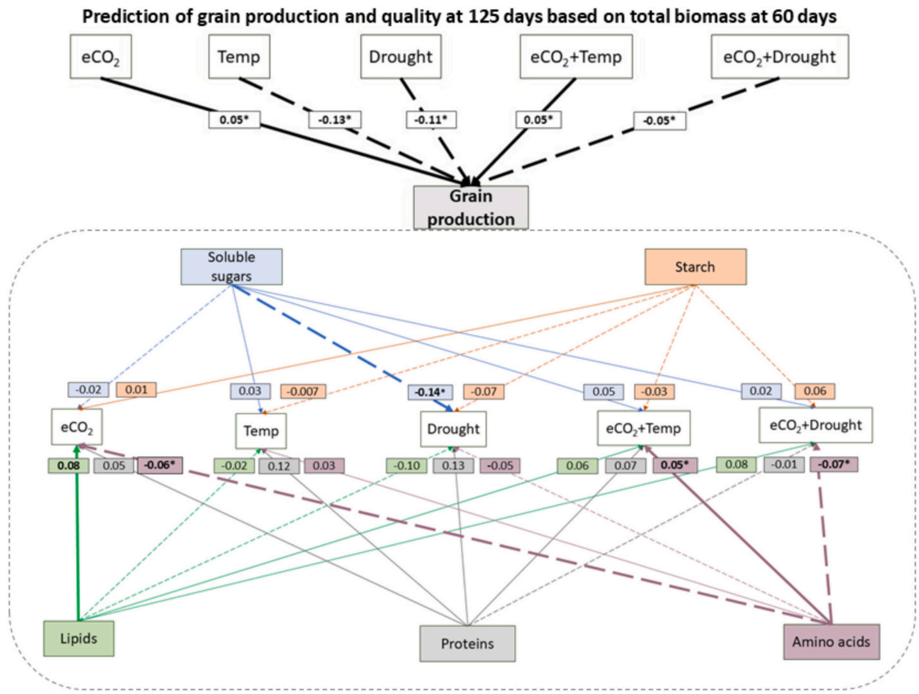


Fig. 6. Path diagram based on Generalized Linear Models (GLM) results shows the relationships between total biomass at 60 DAE under elevated CO_2, temperature, and drought (and combined treatments), along with their effects on soybean production and quality compared to ambient conditions at 125 DAE. The treatments include [aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Drought (400 ppm CO_2 + watering reduction); and eCO_2+Drought (800 ppm CO_2 + watering reduction)]. Solid lines represent positive values, and dashed lines indicate negative values for soybean grain and quality. Bold labels with an asterisk (*) and thick lines highlight statistically significant effects ($P < 0.05$).

Table 2
Generalized Linear Models (GLM) using total biomass at 60 DAE to predict production and quality soybean grain at 125 DAE, according to treatments applied include [aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); aCO_2 (400 ppm CO_2 + ambient temperature); eCO_2 (800 ppm CO_2 + ambient temperature); Drought (400 ppm CO_2 + watering reduction); and eCO_2+Drought (800 ppm + watering reduction)]. Model outputs were used to quantify effects in the path analysis (Fig. 6). Significant effects P-values < 0.05 are in bold.

Treatments	eCO_2	Temp	Drought	eCO_2+Temp	eCO_2+Drought
Grain	0.002	< 0.001	0.011	0.003	0.046
Soluble sugars	0.322	0.615	0.046	0.089	0.631
Starch	0.687	0.933	0.471	0.424	0.304
Lipids	0.017	0.764	0.234	0.097	0.143
Proteins	0.125	0.129	0.128	0.059	0.783
Amino acids	0.007	0.563	0.322	0.029	0.048

Table 3
Predictive machine learning models (XGBoost and CatBoost) comparison prediction for soybean grain production and quality under Triple Effect using experimental data with elevated CO_2 + Temperature and elevated CO_2 + Drought. The performance of the models was evaluated by the root mean square error (RMSE).

	XGBoost		CatBoost	
Triple Effect	Predicted	RMSE	Predicted	RMSE
Production	10.46 ± 1.65	0.04	10.51 ± 1.63	1.50
Soluble sugars	45.53 ± 2.08	0.14	48.07 ± 2.02	9.20
Starch	16.71 ± 0.91	0.08	16.72 ± 0.83	0.96
Lipids	25.54 ± 0.40	0.05	25.56 ± 0.31	1.69
Proteins	35.79 ± 0.55	0.11	35.49 ± 0.51	0.90
Amino acids	158.65 ± 23.45	16.6	178.17 ± 23.19	34.99

significant change (Myers et al., 2014; Palacios et al., 2019). These differences suggest that protein response to elevated CO_2 depends heavily on context and may vary with environmental conditions, soil nitrogen levels, cultivar type, and developmental stage during grain filling (Taub et al., 2008). Elevated CO_2 increased the concentration of soluble sugars and starch, but did not significantly change total protein levels (Figs. 3–5). This shift reflects a metabolic reallocation in which enhanced carbon assimilation promotes carbohydrate and lipid accumulation. Similar responses are seen in other C_3 crops, where elevated CO_2 reduces photorespiration and alters nitrogen assimilation (Ainsworth & Long, 2021; Boote et al., 1997; Myers et al., 2014). These compositional changes have important implications for soybean nutritional value, food quality, and industrial processing (Myers et al., 2014). Increased lipid content may enhance oil yield and improve industrial extraction efficiency (Singh et al., 2016; Soares et al., 2019). However, increased carbohydrate concentrations may influence functional and technological properties, including water absorption capacity, emulsification, texture, and protein functionality, which are critical for soybean derived food products such as tofu, soy milk, and protein isolates (Guan et al., 2021; Myers et al., 2014).

Under climate change scenarios, interactions between elevated CO_2 and temperature and drought may further influence grain composition by altering carbon allocation, nitrogen assimilation, and metabolic regulation during seed filling (Gray et al., 2016; Jin et al., 2017; Palacios et al., 2019). For example, eCO_2 may partially mitigate drought–induced reductions in yield through improving WUE (Li et al., 2024). While eCO_2 may increase oil concentration in grain but main C/N (Palacios et al., 2019). These shifts highlight an important trade–off between increased productivity and potential changes in nutritional quality. Therefore, understanding how climate change factors affect both yield and grain composition is necessary for developing soybean cultivars and management strategies that maintain nutritional quality under environmental conditions.

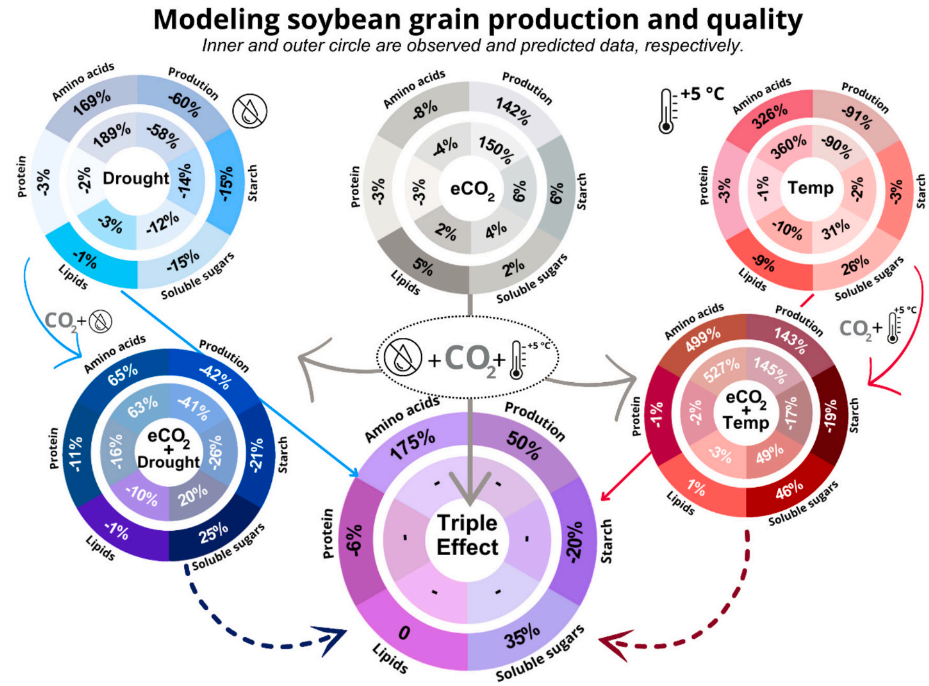


Fig. 7. Modelling soybean grain production under combined stresses. Percentage of response changes related to ambient treatment in observed values (inner circle) and model-predicted values (outer circle) for each treatment: [eCO_2 (800 ppm CO_2 + ambient temperature); Temp (400 ppm CO_2 + 5 °C); eCO_2+Temp (800 ppm CO_2 + 5 °C); Drought (400 ppm CO_2 + watering reduction); eCO_2+Drought (800 ppm CO_2 + watering reduction); and predicted Triple Effect (eCO_2+Temp * eCO_2+Drought)]. Values show relative changes compared to the ambient treatment for grain production and quality (soluble sugars, starch, lipids, proteins, and amino acids). XGBoost was used for the Triple Effect predition, while Generalized Linear Models (GLMs) were used for simple and dual treatments.

4.4. Temperature and drought reduce grain production

Rising global temperatures and more frequent droughts pose significant threats to agricultural productivity worldwide, with soybeans being especially vulnerable (Long, 2025). High temperatures speed up developmental stages, shorten the grain–filling period, and decrease pollen viability and grain formation, resulting in substantial yield losses (Müller & Rieu, 2016). In this study, high–temperature stress caused a 91% reduction in grain biomass (Fig. 7), highlighting the critical sensitivity of soybean reproductive processes to heat stress.

Drought also caused significant reductions in grain production, consistent with previous reports that water limitation reduces pod formation and accelerates senescence (Ferris et al., 1999; Gray et al., 2016; Jin et al., 2017, 2018; Singh et al., 2021). The observed 60% decline in grain yield under Drought (Fig. 7) matches earlier studies showing significant drops in seed biomass during water shortages (A. Wang, et al., 2018). When combined with eCO_2, the negative effects of drought were partly lessened, leading to a 42% reduction in yield. This reduction reflects improved water–use efficiency due to decreased stomatal conductance usually seen with elevated CO_2 (Singh et al., 2021), which lowers transpiration and helps maintain productivity during soil–water shortages. Although temperature and drought alone limit grain formation, elevated CO_2 is physiologically beneficial by increasing carbon assimilation and hydraulic efficiency. However, these benefits are not enough to fully counteract the yield losses caused by extreme heat or extended drought. This underscores the urgent need for integrated modeling, adaptive management, and breeding strategies to ensure soybean productivity and grain quality amid increasing climate stresses.

4.5. Temperature coupled with elevated CO_2 promotes grain biomass

The interaction between eCO_2 and Tempshowed compensatory effects on soybean grain biomass and quality (Fig. 7). The mitigating influence of eCO_2 under high temperature stress can be linked to enhanced photosynthesis, improved water–use efficiency, and changes in carbohydrate partitioning (Habermann et al., 2019; Haworth et al., 2021). Model projections indicated that eCO_2+Temp increased grain production by about 143% compared to aCO_2, closely matching the observed 145% increase. Elevated CO_2 can partly counteract productivity losses caused by temperature increases, especially under moderate warming scenarios (Ainsworth & Long, 2021; Gray et al., 2016; Jin et al., 2017; Singh et al., 2021). However, this synergistic effect also caused changes in grain biochemistry. The eCO_2+Temp treatment altered carbohydrate profiles, likely due to shifts in carbon allocation and respiratory activity during seed development. These metabolic changes may disturb the source–sink balance, leading to higher respiratory demand and less starch accumulation (Yang et al., 2023). Elevated CO_2 may help buffer the negative effects of moderate temperature increases on yield, but related shifts in carbon metabolism can affect grain quality.

4.6. The Triple Effect may increase grain production, but decrease food quality

The machine learning projections of soybean production and quality grain performance to Triple Effect using experimental data combined high temperature under elevated CO_2 and drought under elevated CO_2. XGBoost showed superior predictive accuracy for both soybean grain yield and biochemical composition under the Triple Effect (RMSE = 0.04) compared with CatBoost (RMSE = 1.50) (Table 3). XGBoost often equals or outperforms neural–network approaches in capturing nonlinear and multivariate stress interactions (Bhat et al., 2024; Huber et al., 2022; M'hamdi et al., 2024; Taniushkina et al., 2024). Model predictions and experimental data indicated an average 50% increase in grain production under combined climate treatments, reflecting the compensatory effect of eCO_2 on Temp and Drought–induced yield losses (Fig. 7). The predicted increase in grain yield under the Triple Effect aligns partially with evidence that elevated CO_2 can stimulate photosynthetic carbon assimilation and biomass accumulation in soybean, often resulting in yield gains under moderate stress conditions

(Ainsworth & Long, 2021). However, long-term field experiments have demonstrated that severe high temperatures and drought can reduce or eliminate the CO_2 fertilization effect (Gray et al., 2016). In the Triple Effect prediction, grain quality showed a 35% rise in soluble sugars and a 175% increase in amino acid content, along with declines of 20% in starch and 6% in protein (Fig. 7). High amino acid accumulation is typical of stress adaptation, serving as an osmoprotective and redox–buffering mechanism (Song et al., 2025). While this response may improve short–term tolerance, it also indicates a metabolic reprogramming that emphasizes immediate stress mitigation over carbon storage in proteins and polysaccharides. The projected reductions in starch and protein content are consistent with reports that elevated CO_2 frequently alters carbon–nitrogen balance, often leading to reduced protein concentration due to nitrogen dilution effects (Myers et al., 2014; Taub et al., 2008). The stability of lipid concentration observed in our projections also agrees with previous findings indicating that oil content in soybean tends to be less sensitive to CO_2 enrichment compared to protein (Palacios et al., 2019; Taub et al., 2008). The predicted increase in soluble sugars and total amino acids may reflect stress-induced metabolic reprogramming, as combined heat and drought can promote osmolyte accumulation and shifts in primary metabolism to support cellular protection (Mahalingam, 2014; Pascual et al., 2022). Experimental studies demonstrate that combined stressors can trigger emergent physiological responses that are not predictable from individual stress effects (Gray et al., 2016; Pascual et al., 2022). Consequently, the performance of XGBoost suggests that ensemble learning approaches are effective in capturing complex relationships (Chen & Guestrin, 2016).

In a recent meta-analysis on herbaceous species exposed to three factor combinations, Yoldi-Achalandabaso et al. (2025) highlighted the scarcity of factorial studies and identified soybean as being represented by only one trifactorial experiment, conducted by Singh et al. (2021), without quality analysis in grain. In that study, when compared to the control conditions (ambient CO_2, optimal temperature, and well-watered), the combined stress treatment resulted in an overall 45% reduction in seed production, reinforcing the dominant constraining role of water limitation under multifactorial stress. Seed production tended to be greater under high temperature compared to the optimum regime when elevated CO_2 was supplied with water stress (Singh et al., 2021), suggesting that moderate warming may enhance carbon assimilation when water is not limited. However, the three interactions were not consistently significant across all evaluated traits, indicating that the outcome of the three factors depends strongly on stress intensity and environmental context. Our XGBoost projections under the Triple Effect scenario predicted a 50% increase in grain production relative to ambient conditions, accompanied by quality grain prediction with a 35% rise in soluble sugars and a 175% increase in total amino acids, reductions of approximately 20% in starch and 6% in protein. The divergence in magnitude and direction of production responses between our predictive model and the trifactorial experiment may reflect differences in OTCs and SPAR (Soil Plant Atmosphere Research chambers) - used by Singh et al. (2021) to stress treatments. Jin et al. (2017), for example, demonstrated that drought remains the dominant limiting factor for soybean production under future climate scenarios, even when CO_2 fertilization effects are considered, and that concurrent stresses can amplify production losses through interactions. Nevertheless, consistent with both the three-factor experiment and large-scale process-based APSIM modeling studies (Jin et al., 2017), our results should be interpreted as scenario-based projections rather than direct empirical equivalents, especially under severe or field-realistic stress combinations where drought may overestimate CO_2 benefits. The ability to quantify these yield–quality trade–offs through rapid, high–precision modeling offers valuable insights for breeding, crop management, and climate policy. From a food system perspective, the observed reductions in starch and protein levels are potential vulnerabilities in the nutritional quality of soybean grains under future climate scenarios. Therefore, while elevated CO_2 may sustain or even increase grain yield under combined stresses, its benefits are offset by changes in composition that could weaken feed and food value.

4.7. Triple Effect compared with interactions under field conditions

A controlled environment cannot fully reproduce the spatial and temporal heterogeneity of field conditions, where crops are simultaneously influenced by fluctuating radiation, soil variability, vapor pressure deficit, pest pressure, nutrient dynamics, and extreme climate events (Langstroff et al., 2022). Large-scale modeling and field-based assessments indicate that elevated CO_2 often only partially offsets heat, and drought impacts under realistic climate variability (Jin et al., 2017; Yang et al., 2026). Moreover, climate change may amplify environmental feedback, including soil degradation and nutrient losses, further modifying crop responses beyond chamber-based observations (Singh et al., 2021). Interactions observed under controlled systems may overestimate compensatory CO_2 effects or underestimate the complexity of field-scale responses, where cycle duration, canopy temperature, soil water dynamics, and energy balance feedback interact (Battisti et al., 2017). On the other hand, Kothari et al. (2022) demonstrate the advantages of employing an ensemble of grain legume models to improve the robustness of climate change projections for food production. At the same time, the results highlight the need for continued refinement of soybean models, particularly through experimental data generated under elevated CO_2 and temperature conditions.

Process-based models simulate physiological processes and allow mechanistic partitioning of stress effects (Battisti et al., 2017; Jin et al., 2017). For example, APSIM simulations indicate that elevated CO_2 only partially offsets drought-induced soybean yield losses, with diminishing fertilization benefits under increasing drought frequency (Jin et al., 2017). In contrast, ensemble machine learning models can flexibly learn from multivariate data. Nonetheless, purely data-driven approaches may suffer from limited extrapolation capacity and reduced interpretability outside the training domain, particularly under novel climate combinations (Yang et al., 2026). Our analysis was conducted using a single soybean cultivar under controlled chamber conditions, which constrains generalizability across genotypes and field environments. Large-scale climate impact assessments show substantial variation in stress sensitivity across regions, management systems, and cultivars (Jin et al., 2017), and warming has been shown to differentially affect cropping systems depending on local adaptation capacity (Jin et al., 2017, 2018). Such hybrid approaches may ultimately bridge the strengths of mechanistic crop models and advanced ensemble learning (Yang et al., 2026), improving predictive confidence for multifactor climate stress impacts on crop productivity and quality.

4.8. Limitation of triple effect predictions

Extrapolating soybean responses from dual-stress experiments (e.g., $eCO_2 \times$ Temp or $eCO_2 \times$ Drought) to Triple Effect scenarios introduces uncertainty due to not experimental validation. Plants often involve synergistic or antagonistic effects mediated by complex physiological trade-offs among carbon assimilation, water relations, respiration, and reproductive development (Mahalingam, 2014; Pascual et al., 2022). For example, while elevated CO_2 can enhance photosynthesis gain and WUE under moderate drought, concurrent high temperatures may increase photorespiration and respiratory carbon losses, potentially negating the CO_2 fertilization effect under severe stress (Gray et al., 2016; Webber et al., 2022). Moreover, stress intensity and developmental stage can substantially alter the direction and magnitude of responses, making model extrapolation beyond experimentally tested combinations inherently uncertain (Wang & Liu, 2021). Process-based and statistical models calibrated under single or dual stress conditions may fail to capture emergent thresholds or tipping points that arise under three stress exposures, particularly when feedback mechanisms and metabolic reprogramming are involved (Webber et al., 2022). The projections should be interpreted as scenario-based predictions that require further validation under fully integrated Triple Effect experimental conditions.

The limitation of the machine learning framework used in this study is the risk of overfitting when modeling complex interactions among the Triple Effect. However, a 5-fold cross-validation was performed to minimize overfitting risk (Table S4). Ensemble algorithms such as XGBoost are highly flexible and can capture intricate multivariate patterns, but this flexibility may lead to overly optimistic performance if the model fits noise rather than biologically meaningful signals, particularly when training datasets are limited or derived from controlled conditions (Chen & Guestrin, 2016). Although elevated CO_2 can enhance carbon assimilation and modify carbohydrate partitioning in C3 crops, these responses depend strongly on sink strength, nutrient availability, and stress intensity (Ainsworth & Long, 2021). Models may amplify statistical associations that do not generalize to broader field conditions. Thus, despite applying cross-validation, the Triple Effect projections should be interpreted as scenario-based inferences requiring validation with independent multifactorial and field-scale datasets.

5. Conclusion

This study aimed to achieve two related objectives with direct implications for climate–resilient soybean production. First, we showed that generalized linear models (GLMs) based on early–stage physiological and morphological traits—especially total biomass at 60 DAE—can accurately predict grain yield and compositional quality at 125 DAE across various environmental conditions. The low prediction errors and the strong agreement between observed and predicted values highlight the potential of GLMs as effective tools for forecasting crop performance. The use of early–stage biomass as a predictive indicator adds value by enabling better agronomic management and guiding breeding programs toward improved stress tolerance. Second, by modeling the combined effects of eCO_2, high temperature, and drought (Triple Effect), we revealed how soybean responses are influenced by concurrent climate stressors. Combining experimental data from dual–factor treatments with advanced machine learning methods, XGBoost consistently outperformed CatBoost in prediction accuracy, effectively capturing nonlinear and interactive relationships among stress variables. Under the Triple Effect model, projections showed a 50% increase in grain production, along with higher levels of soluble sugars (+35%) and amino acids (+175%), but with significant declines in starch (−20%) and protein (−6%). Elevated CO_2 may partially offset yield losses caused by temperature and drought, but at the expense of nutritional quality. This highlights the urgent need for integrated modeling frameworks and breeding programs that jointly improve yield stability, nutritional value, and stress resilience. Ultimately, combining empirical experiments with predictive modeling offers a powerful approach to anticipate climate–-driven changes in crop productivity and to develop adaptive strategies for sustainable food systems.

CRediT authorship contribution statement

Janaina da Silva Fortirer: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Adriana Grandis:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Carmen Eusebia Palacios Jara:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Débora Pagliuso:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Leandro Francisco de Oliveira:** Writing – review & editing, Writing – original draft, Methodology, Data curation. **Eveline Queiroz de Pinho Tavares:** Writing – review & editing, Methodology. **Lauana Pereira de Oliveira:** Writing – review & editing, Methodology. **Plínio B. Camargo:** Writing – review & editing, Methodology. **Eny Iochevet Segal Floh:** Writing – review & editing, Methodology. **Cibele M. Russo:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis. **Marcos S. Buckeridge:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank CAPES for the fellowships granted and Dr. Vanoli Fronza, a researcher at the Embrapa Soybean Division (Londrina, PR), for providing the seeds of the cultivars used. This work has been partially financed by the National Institute of Science and Technology of Bioethanol (INCT−Bioethanol – FAPESP 2008/57908-6 and CNPq 574002/2008−1). JSF is financed by the FAPESP (2022/14886-0) and (2025/15362-2). LFO (RCGI 371055 and FAPESP 2024/12357−5). MSB are fellow researchers of CNPq (National Council for Scientific and Technological Development, Brazil).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2026.119004.

Data availability

Data will be made available on request.

References

Ainsworth, E. A., & Lemonnier, P. (2018). Phloem function: A key to understanding and manipulating plant responses to rising atmospheric [CO_2]? *Current Opinion in Plant Biology, 43*, 50–56. https://doi.org/10.1016/j.pbi.2017.12.003

Ainsworth, E. A., & Long, S. P. (2021). 30 years of free-air carbon dioxide enrichment (FACE): What have we learned about future crop productivity and its potential for adaptation? *Global Change Biology, 27*(1), 27–49. https://doi.org/10.1111/gcb.15375

Amaral, L. I. V., do Gaspar, M., Costa, P. M. F., Aidar, M. P. M., & Buckeridge, M. S. (2007). Novo método enzimático rápido e sensível de extração e dosagem de amido em materiais vegetais. *Hoehnea, 34*, 425–431. https://doi.org/10.1590/S2236-89062007000400001

ANEC. (2022). *Associação Nacional dos Exportadores de Cereais*. In Retrieved from *https://anec.com.br/*. Accessed February 27, 2026.

Arenque, B. C., Grandis, A., Pocius, O., de Souza, A. P., & Buckeridge, M. S. (2014). Responses of Senna reticulata, a legume tree from the Amazonian floodplains, to elevated atmospheric CO_2 concentration and waterlogging. *Trees, 28*, 1021–1034. https://doi.org/10.1007/s00468-014-1015-0

Assad, E. D., & Assad, M. L. R. C. L. (2024). Mudanças do clima e agropecuária: impactos, mitigação e adaptação. *Desafios e oportunidades. Estudos Avançados, 38*, 271–292. https://doi.org/10.1590/s0103-4014.202438112.015

Banga, S. S., & Kang, M. S. (2014). Developing climate-resilient crops. *Journal of Crop Improvement, 28*(1), 57–87. https://doi.org/10.1080/15427528.2014.865410

Battisti, R., Sentelhas, P. C., & Boote, K. J. (2017). Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. *Field Crops Research, 200*, 28–37. https://doi.org/10.1016/j.fcr.2016.10.004

Beck, J. (2013). Predicting climate change effects on agriculture from ecological niche modeling: Who profits, who loses? *Climatic Change, 116*(2), 177–189. https://doi.org/10.1007/s10584-012-0481-x

Bhat, S. A., Qadri, S. A. A., Dubbey, V., Sofi, I. B., & Huang, N.-F. (2024). Impact of crop management practices on maize yield: Insights from farming in tropical regions and predictive modeling using machine learning. *Journal of Agriculture and Food Research, 18*, Article 101392. https://doi.org/10.1016/j.jafr.2024.101392

Boote, K., Pickering, N., & Allen, L., Jr. (1997). Plant modeling: Advances and gaps in our capability to predict future crop growth and yield in response to global climate change. *Advances in Carbon Dioxide Effects Research, 179–228*. https://doi.org/10.2134/asaspecpub61.c10

Bui, Q.-T., Pham, Q.-T., Pham, V.-M., Tran, V.-T., Nguyen, D.-H., Nguyen, Q.-H., … Vu, V.-M. (2024). Hybrid machine learning models for aboveground biomass estimations. *Ecological Informatics, 79*, Article 102421. https://doi.org/10.1016/j.ecoinf.2023.102421

Cai, Y., Chen, L., Zhang, Y., Yuan, S., Su, Q., Sun, S., Wu, C., Yao, W., Han, T., & Hou, W. (2020). Target base editing in soybean using a modified CRISPR/Cas9 system. *Plant Biotechnology Journal, 18*(10), 1996. https://doi.org/10.1111/pbi.13386

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Chenu, K., Porter, J. R., Martre, P., Basso, B., Chapman, S. C., Ewert, F., … Asseng, S. (2017). Contribution of crop models to adaptation in wheat. *Trends in Plant Science, 22*(6), 472–490. https://doi.org/10.1016/j.tplants.2017.02.003

Christie, W. W., & Hutton, J. (1993). Preparation of ester derivatives of fatty acids for chromatographic analysis. *Advances in Lipid Methodology, 2*(69), Article e111.

CONAB. (2021). *Companhia Nacional de Abastecimento*. Retrieved from https://www.gov.br/conab/pt-br. Accessed February 27, 2026.

DaMatta, F. M., Grandis, A., Arenque, B. C., & Buckeridge, M. S. (2010). Impacts of climate changes on crop physiology and food quality. *Food Research International, 43*(7), 1814–1823. https://doi.org/10.1016/j.foodres.2009.11.001

De Winter, J. C., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological Methods, 21*(3), 273. https://doi.org/10.1037/met0000079.supp

Di Paola, A., Di Giuseppe, E., Gutierrez, A. P., Ponti, L., & Pasqui, M. (2023). Climate stressors modulate interannual olive yield at province level in Italy: A composite index approach to support crop management. *Journal of Agronomy and Crop Science, 209*(4), 475–488. https://doi.org/10.1111/jac.12636

Digrado, A., Montes, C. M., Baxter, I., & Ainsworth, E. A. (2024). Seed quality under elevated CO_2 differs in soybean cultivars with contrasting yield responses. *Global Change Biology, 30*(2), Article e17170. https://doi.org/10.1111/gcb.17170

Ekele, J. U., Webster, R., Perez de Heredia, F., Lane, K. E., Fadel, A., & Symonds, R. C. (2025). Current impacts of elevated CO_2 on crop nutritional quality: a review using wheat as a case study. *Stress Biology, 5*(1), 34. https://doi.org/10.1007/s44154-025-00217-w

Ewert, F., Porter, J. R., & Rounsevell, M. D. (2007). Crop models, CO_2, and climate change. *Science, 26;315(5811):459-60*. https://doi.org/10.1126/science.315.5811.459c

FAO. (2023). *Food and Agriculture Organization of the United Nations Crops*. In Retrieved from http://www.fao.org/faostat/en/#data/QC/visualize. Accessed February 27, 2026.

Ferris, R., Wheeler, T., Ellis, R., & Hadley, P. (1999). Seed yield after environmental stress in soybean grown under elevated CO_2. *Crop Science, 39*(3), 710–718. https://doi.org/10.2135/cropsci1999.0011183X003900030018x

Fortirer, J. S., Grandis, A., Pagliuso, D., de Castanho, C. T., & Buckeridge, M. S. (2023). Meta-analysis of the responses of tree and herb to elevated CO_2 in Brazil. *Scientific Reports, 13*(1). https://doi.org/10.1038/s41598-023-40783-5

Fu, Z., Niu, S., & Dukes, J. S. (2015). What have we learned from global change manipulative experiments in China? *A meta-analysis. Scientific Reports, 5*(19), 1–11. https://doi.org/10.1038/srep12344

Gong, Z., Zhang, X., Zong, Y., Zhang, D., Shi, X., Hao, X., & Li, P. (2025). Co-elevated CO_2 concentration and temperature enhance the carbon assimilation and lipid metabolism in a high-oil soybean (*Glycine max* (L.) Merr.) variety. *Plant Physiology and Biochemistry*. https://doi.org/10.1016/j.plaphy.2025.110243

Gray, S. B., Dermody, O., Klein, S. P., Locke, A. M., Mcgrath, J. M., Paul, R. E., Rosenthal, D. M., Ruiz-Vera, U. M., Siebers, M. H., Strellner, R., & others. (2016). Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean. Nature Plants, 2(9), 1–8. doi:https://doi.org/10.1038/nplants.2016.132.

Guan, X., Zhong, X., Lu, Y., Du, X., Jia, R., Li, H., & Zhang, M. (2021). Changes of soybean protein during tofu processing. *Foods, 10*(7), 1594. https://doi.org/10.3390/foods10071594

Habermann, E., Dias de Oliveira, E. A., Contin, D. R., San Martin, J. A., Curtarelli, L., Gonzalez-Meler, M. A., & Martinez, C. A. (2019). Stomatal development and conductance of a tropical forage legume are regulated by elevated [CO_2] under moderate warming. *Frontiers in Plant Science, 10*, 609. https://doi.org/10.3389/fpls.2019.00609

Hallgren, K. A., McCabe, C. J., King, K. M., & Atkins, D. C. (2019). Beyond path diagrams: Enhancing applied structural equation modeling research through data visualization. *Addictive Behaviors, 94*, 74–82. https://doi.org/10.1016/j.addbeh.2018.08.030

Hancock, J., & Khoshgoftaar, T. (2020). CatBoost for big data: An interdisciplinary review. *Journal of Big Data, 7*(1), 94. https://doi.org/10.1186/s40537-020-00369-8

Haworth, M., Marino, G., Loreto, F., & Centritto, M. (2021). Integrating stomatal physiology and morphology: Evolution of stomatal control and development of future crops. *Oecologia, 197*(4), 867–883. https://doi.org/10.1007/s00442-021-04857-3

Heinemann, B., & Hildebrandt, T. M. (2021). The role of amino acid metabolism in signaling and metabolic adaptation to stress-induced energy deficiency in plants. *Journal of Experimental Botany, 72*(13), 4634–4645. https://doi.org/10.1093/jxb/erab182

Hoagland, K. D., Rosowski, J. R., Gretz, M. R., & Roemer, S. C. (1993). Diatom extracellular polymeric substances: Function, fine structure, chemistry, and physiology. *Journal of Phycology, 29*(5), 537–566. https://doi.org/10.1111/j.0022-3646.1993.00537.x

Huber, F., Yushchenko, A., Stratmann, B., & Steinhage, V. (2022). Extreme Gradient Boosting for yield estimation compared with Deep Learning approaches. *Computers and Electronics in Agriculture, 202*(107346), 2021. https://doi.org/10.1016/j.compag.2022.107346IPCC

Jardim Amorim, D., Pereira dos Santos, A. R., Nunes da Piedade, G., Quelvia de Faria, R., Amaral da Silva, E. A., & Pereira Sartori, M. M. (2021). The use of the generalized linear model to assess the speed and uniformity of germination of corn and soybean seeds. *Agronomy, 11*(3), 588. https://doi.org/10.3390/agronomy11030588

Jiang, Z., van Zanten, M., & Sasidharan, R. (2025). Mechanisms of plant acclimation to multiple abiotic stresses. *Communications Biology, 8*, 655. https://doi.org/10.1038/s42003-025-08077-w

Jin, Z., Ainsworth, E. A., Leakey, A. D., & Lobell, D. B. (2018). Increasing drought and diminishing benefits of elevated carbon dioxide for soybean yields across the US Midwest. *Global Change Biology, 24*(2), e522–e533. https://doi.org/10.1111/gcb.13946

Jin, Z., Zhuang, Q., Wang, J., Archontoulis, S. V., Zobel, Z., & Kotamarthi, V. R. (2017). The combined and separate impacts of climate extremes on the current and future US rainfed maize and soybean production under elevated CO_2. *Global Change Biology, 23*(7), 2687–2704. https://doi.org/10.1111/gcb.13617

Keenan, T. F., Luo, X., Stocker, B. D., et al. (2023). A constraint on historic growth in global photosynthesis due to rising CO_2. *Nature Climate Change, 13*, 1376–1381. https://doi.org/10.1038/s41558-023-01867-2

Khan, M. H. U., Wang, S., Wang, J., Ahmar, S., Saeed, S., Khan, S. U., … Feng, X. (2022). Applications of artificial intelligence in climate-resilient smart-crop breeding. *International Journal of Molecular Sciences, 23*(19), 11156. https://doi.org/10.3390/ijms231911156

Kohavi, R. & others. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai, 14(2), 1137–1145.

Kothari, K., Battisti, R., Boote, K. J., Archontoulis, S. V., Confalone, A., Constantin, J., & Salmerón, M. (2022). Are soybean models ready for climate change food impact assessments? *European Journal of Agronomy, 135*, Article 126482. https://doi.org/10.1016/j.eja.2022.126482

Langstroff, A., Heuermann, M. C., Stahl, A., & Junker, A. (2022). Opportunities and limits of controlled-environment plant phenotyping for climate response traits. *Theoretical and Applied Genetics, 135*(1), 1–16. https://doi.org/10.1007/s00122-021-03892-1

Leakey, A.D.B., Ainsworth, E.A., Bernacchi, C.J., Rogers, A., Long, S.P., & Ort, D.R. (2009). Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. Journal of Experimental Botany, 60(10), 2859–2876. ttps://doi.org/10.1093/jxb/erp096.

Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature, 529*(7584), 84–87. https://doi.org/10.1038/nature16467

Li, S., Agathokleous, E., Li, S., Xu, Y., Xia, J., & Feng, Z. (2024). Climate gradient and leaf carbon investment influence the effects of climate change on water use efficiency of forests: a meta-analysis. *Plant, Cell & Environment, 47*, 1070–1083. https://doi.org/10.1111/pce.14777

Li, Y., Zeng, H., Zhang, M., Wu, B., Zhao, Y., Yao, X., Cheng, T., Qin, X., & Wu, F. (2023). A county-level soybean yield prediction framework coupled with XGBoost and multidimensional feature engineering. *International Journal of Applied Earth Observation and Geoinformation, 118*, Article 103269. https://doi.org/10.1016/j.jag.2023.103269

Loague, K., & Green, R. E. (1991). Statistical and graphical methods for evaluating solute transport models: Overview and application. *Journal of Contaminant Hydrology, 7*(1–2), 51–73. https://doi.org/10.1016/0169-7722(91)90038-3

Long, S. P. (2025). Needs and opportunities to future-proof crops and the use of crop systems to mitigate atmospheric change. *Philosophical Transactions B, 380*(1927), 20240229. https://doi.org/10.1098/rstb.2024.0229

M'hamdi, O., Takács, S., Palotás, G., Ilahy, R., Helyes, L., & Pék, Z. (2024). A comparative analysis of XGBoost and neural network models for predicting some tomato fruit quality traits from environmental and meteorological data. *Plants, 13*(5), 746. https://doi.org/10.3390/plants13050746

Machado, J., Nunes da Silva, M., Vasconcelos, M. W., & Santos, C. S. (2025). The impact of climate change-induced abiotic stresses on the nutritional quality of legume seeds. *Journal of Experimental Botany, 76*(12). https://doi.org/10.1093/jxb/eraf085

Mahalingam, R. (Ed.). (2014). *Combined stresses in plants: physiological, molecular, and biochemical aspects*. Springer. https://doi.org/10.1007/978-3-319-07899-1.

Marengo, J. A., Torres, R. R., & Alves, L. M. (2017). Drought in Northeast Brazil—Past, present, and future. *Theoretical and Applied Climatology, 129*(3–4), 1189–1200. https://doi.org/10.1007/s00704-016-1840-8

Mariotti, F., Tomé, D., & Mirand, P. P. (2008). Converting nitrogen into protein—Beyond 6.25 and Jones' factors. *Critical Reviews in Food Science and Nutrition, 48*(2), 177–184. https://doi.org/10.1080/10408390701279749

Monteiro, L. A., Ramos, R. M., Battisti, R., Soares, J. R., Oliveira, J. C., Figueiredo, G. K., … Lana, M. A. (2022). Potential use of data-driven models to estimate and predict soybean yields at national scale in Brazil. *International Journal of Plant Production, 16*(4), 691–703. https://doi.org/10.1007/s42106-022-00209-0

Mourtzinis, S., Specht, J. E., Lindsey, L. E., Wiebold, W. J., Ross, J., Nafziger, E. D., Kandel, H. J., Mueller, N., Devillez, P. L., Arriaga, F. J., & others. (2015). Climate-induced reduction in US-wide soybean yields underpinned by region-and in-season-specific responses. Nature Plants, 1(2), 1–4. doi:https://doi.org/10.1038/nplants.2014.26.

Müller, F., & Rieu, I. (2016). Acclimation to high temperature during pollen development. *Plant Reproduction, 29*, 107–118. https://doi.org/10.1007/s00497-016-0282-x

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., & Usui, Y. (2014). Increasing CO_2 threatens human nutrition. *Nature, 510*(7503), 139–142. https://doi.org/10.1038/nature13179

Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 135*(3), 370–384. https://doi.org/10.2307/2344614

Nendel, C., Reckling, M., Debaeke, P., Schulz, S., Berg-Mohnicke, M., Constantin, J., Fronzek, S., Hoffmann, M., Jakšić, S., Kersebaum, K.-C., & others. (2023). Future area expansion outweighs increasing drought risk for soybean in Europe. Global Change Biology, 29(5), 1340–1358. doi:https://doi.org/10.1111/gcb.16562.

Nishat, M. H., Khan, M. H. R. B., Ahmed, T., Hossain, S. N., Ahsan, A., El-Sergany, M., … Alresheedi, M. T. (2025). Comparative analysis of machine learning models for predicting water quality index in Dhaka's rivers of Bangladesh. *Environmental Sciences Europe, 37*(1), 31. https://doi.org/10.1186/s12302-025-01078-w

Onat, B., Bakal, H., Gulluoglu, L., & Arıoglu, H. (2017). The effects of high temperature at the growing period on yield and yield components of soybean [*Glycine max* (L.) Merr] varieties. *Turkish Journal of Field. Crops, 22*(2), 178–186. https://doi.org/10.17557/tjfc.356210

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., & Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change, 11*(4), 306–312. https://doi.org/10.1038/s41558-021-01000-1

Palacios, C., Grandis, A., Carvalho, V., Salatino, A., & Buckeridge, M. (2019). Isolated and combined effects of elevated CO_2 and high temperature on the whole-plant biomass and the chemical composition of soybean seeds. *Food Chemistry, 275*, 610–617. https://doi.org/10.1016/j.foodchem.2018.09.052

Pascual, L. S., Segarra-Medina, C., Gómez-Cadenas, A., López-Climent, M. F., Vives-Peris, V., & Zandalinas, S. I. (2022). Climate change-associated multifactorial stress combination: A present challenge for our ecosystems. *Journal of Plant Physiology, 276*, Article 153764. https://doi.org/10.1016/j.jplph.2022.153764

Pasquel, D., Roux, S., Richetti, J., Cammarano, D., Tisseyre, B., & Taylor, J. A. (2022). A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agriculture, 23*(4), 1489–1513. https://doi.org/10.1007/s11119-022-09885-4

Peláez-Vico, M.Á., Sinha, R., Induri, S. P., Lyu, Z., Venigalla, S. D., Vasireddy, D., & Mittler, R. (2024). The impact of multifactorial stress combination on reproductive tissues and grain yield of a crop plant. *The Plant Journal, 117*(6), 1728–1745. https://doi.org/10.1111/tpj.16570

Poudel, S., Vennam, R. R., Shrestha, A., Reddy, K. R., Wijewardane, N. K., Reddy, K. N., & Bheemanahalli, R. (2023). Resilience of soybean cultivars to drought stress during flowering and early-seed setting stages. *Scientific Reports, 13*(1), 1277. https://doi.org/10.1038/s41598-023-28354-0

Prechelt, L. (1998). Early Stopping - But When? In G. B. Orr, & K. R. Müller (Eds.), *Lecture Notes in Computer Science: 1524*. *Neural Networks: Tricks of the Trade*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-49430-8_3.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems, 31*. https://doi.org/10.48550/arXiv.1706.09516

Qiao, Y., Miao, S., Li, Q., Jin, J., Luo, X., & Tang, C. (2019). Elevated CO_2 and temperature increase grain oil concentration but their impacts on grain yield differ between soybean and maize grown in a temperate region. *Science of the Total Environment, 666*, 405–413. https://doi.org/10.1016/j.scitotenv.2019.02.149

Ruiz-Vera, U. M., Siebers, M., Gray, S. B., Drag, D. W., Rosenthal, D. M., Kimball, B. A., & Bernacchi, C. J. (2013). Global warming can negate the expected CO_2 stimulation in photosynthesis and productivity for soybean grown in the Midwestern United States. *Plant Physiology, 162*(1), 410–423. https://doi.org/10.1104/pp.112.211938

Sentelhas, P. C., Battisti, R., de Câmara, G. M. S., Farias, J., Hampf, A., & Nendel, C. (2015). The soybean yield gap in Brazil–magnitude, causes and possible solutions for sustainable production. *The Journal of Agricultural Science, 153*(8), 1394–1411. https://doi.org/10.1017/S0021859615000313

Shanker, A. K., Gunnapaneni, D., Bhanu, D., Vanaja, M., Lakshmi, N. J., Yadav, S. K., … Singh, V. K. (2022). Elevated CO_2 and water stress in combination in plants: Brothers in arms or partners in crime? *Biology, 11*(9), 1330. https://doi.org/10.3390/biology11091330

Shastry, A., Sanjay, H., & Bhanusree, E. (2017). Prediction of crop yield using regression techniques. *International Journal of Soft Computing, 12*(2), 96–102.

Singh, S. K., Barnaby, J. Y., Reddy, V. R., & Sicher, R. C. (2016). Varying response of the concentration and yield of soybean seed mineral elements, carbohydrates, organic acids, amino acids, protein, and oil to phosphorus starvation and CO_2 enrichment. *Frontiers in Plant Science, 7*, 1967. https://doi.org/10.3389/fpls.2016.01967

Singh, S. K., Reddy, V. R., Devi, M. J., & Timlin, D. J. (2021). Impact of water stress under ambient and elevated carbon dioxide across three temperature regimes on soybean canopy gas exchange and productivity. *Scientific Reports, 11*(1), 16511. https://doi.org/10.1038/s41598-021-96037-9

Soares, J., Deuchande, T., Valente, L. M., Pintado, M., & Vasconcelos, M. W. (2019). Growth and nutritional responses of bean and soybean genotypes to elevated CO_2 in a controlled environment. *Plants, 8*(11), 465. https://doi.org/10.3390/plants8110465

Song, Y., John Martin, J. J., Liu, X., Li, X., Hou, M., Zhang, R., … Cao, H. (2025). Unraveling the response of secondary metabolites to cold tolerance in oil palm by integration of physiology and metabolomic analyses. *BMC Plant Biology, 25*(1), 279. https://doi.org/10.1186/s12870-025-06292-5

SoyStats. (2023). *Soy Highlights*. Retrieved from http://soystats.com. Accessed February 27, 2026.

Streiner, D. L. (2005). Finding our way: An introduction to path analysis. *The Canadian Journal of Psychiatry, 50*(2), 115–122. https://doi.org/10.1177/070674370505000207

Taniushkina, D., Lukashevich, A., Shevchenko, V., Belalov, I. S., Sotiriadi, N., Narozhnaia, V., Kovalev, K., Krenke, A., Lazarichev, N., Bulkin, A., & others. (2024). Case study on climate change effects and food security in Southeast Asia. Scientific Reports, 14(1), 16150. doi:https://doi.org/10.1038/s41598-024-65140-y.

Taub, D. R., Miller, B., & Allen, H. (2008). Effects of elevated CO_2 on the protein concentration of food crops: a meta-analysis. *Global Change Biology, 14*(3), 565–575. https://doi.org/10.1111/j.1365-2486.2007.01511.x

Team, R. C. (2024). *R: A language and environment for statistical computing Version 4. 4.3*. Vienna, Austria: R Foundation for Statistical Computing.

Thenveettil, N., Bheemanahalli, R., Reddy, K. N., Gao, W., & Reddy, K. R. (2024). Temperature and elevated CO_2 alter soybean seed yield and quality, exhibiting transgenerational effects on seedling emergence and vigor. *Frontiers in Plant Science, 15*. https://doi.org/10.3389/fpls.2024.1427086

USDA. (2024). *Cropexplorer [Dataset]*. Retrieved from https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=2222000. Accessed February 27, 2026.

Van Eeuwijk, F. A., Bustos-Korts, D., Millet, E. J., Boer, M. P., Kruijer, W., Thompson, A., Malosetti, M., Iwata, H., Quiroz, R., Kuppe, C., & others. (2019). Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. Plant Science, 282, 23–39. doi:https://doi.org/10.1016/j.plantsci.2018.06.018.

Wang, A., Lam, S. K., Hao, X., Li, F. Y., Zong, Y., Wang, H., & Li, P. (2018). Elevated CO_2 reduces the adverse effects of drought stress on a high-yielding soybean (*Glycine max* (L.) Merr.) cultivar by increasing water use efficiency. *Plant Physiology and Biochemistry, 132*, 660–665. https://doi.org/10.1016/j.plaphy.2018.10.016

Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R. P., & Asseng, S. (2017). The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nature Plants, 3*(8), 17102. https://doi.org/10.1038/nplants.2017.102

Wang, N., Wang, E., Wang, J., Zhang, J., Zheng, B., Huang, Y., & Tan, M. (2018). Modelling maize phenology, biomass growth and yield under contrasting temperature conditions. *Agricultural and Forest Meteorology, 250*, 319–329. https://doi.org/10.1016/j.agrformet.2018.01.005

Wang, X., & Liu, F. (2021). Effects of Elevated CO_2 and Heat on Wheat Grain Quality. *Plants, 10*(5), 1027. https://doi.org/10.3390/plants10051027

Wang, Y., Zhang, Q., Yu, F., Zhang, N., Zhang, X., Li, Y., Wang, M., & Zhang, J. (2024). Progress in Research on Deep Learning-Based Crop Yield Prediction. *Agronomy, 14*(10), 2264. https://doi.org/10.3390/agronomy14102264

Webber, H., Rezaei, E. E., Ryo, M., & Ewert, F. (2022). Framework to guide modeling single and multiple abiotic stresses in arable crops. *Agriculture, Ecosystems & Environment, 340*, Article 108179. https://doi.org/10.1016/j.agee.2022.108179

Wickham, H., Francois, R., & Wickham, M. H. (2014). *Package 'dplyr'*.

Wu, X., Liu, H., Li, X., Ciais, P., Babst, F., Guo, W., Zhang, C., Magliulo, V., Pavelka, M., Liu, S., Huang, Y., Wang, P., Shi, C., & Ma, Y. (2018). Differentiating drought legacy effects on vegetation growth over the temperate Northern Hemisphere. *Global Change Biology, 24*(1), 504–516. https://doi.org/10.1111/gcb.13920

Wu, Y., Wang, E., He, D., Liu, X., Archontoulis, S. V., Huth, N. I., … Yang, W. (2019). Combine observational data and modelling to quantify cultivar differences of soybean. *European Journal of Agronomy, 111*, Article 125940. https://doi.org/10.1016/j.eja.2019.125940

Yang, J., Liu, L., Yang, Q., Jia, X., Peng, B., Guan, K., & Jin, Z. (2026). Knowledge-guided graph machine learning improves corn yield mapping in the US Midwest. *Remote Sensing of Environment, 335*, Article 115287. https://doi.org/10.1016/j.rse.2026.115287

Yang, L., Song, W., Xu, C., Sapey, E., Jiang, D., & Wu, C. (2023). Effects of high night temperature on soybean yield and compositions. *Frontiers in Plant Science, 14*, 1065604. https://doi.org/10.3389/fpls.2023.1065604

Yoldi-Achalandabaso, A., Vicente, R., Muñoz-Rueda, A., & Pérez-López, U. (2025). Unveiling herbaceous plant responses to future triple interaction of drought and elevated temperature and [CO_2]: Systematic review and meta-analysis. *Physiologia Plantarum, 177*(1), Article e70125. https://doi.org/10.1111/ppl.70125

Zandalinas, S. I., & Mittler, R. (2022). Plant responses to multifactorial stress combination. *New Phytologist, 234*(4), 1161–1167. https://doi.org/10.1111/nph.18087

Zandalinas, S. I., Peláez-Vico, M.Á., Sinha, R., Pascual, L. S., & Mittler, R. (2024). The impact of multifactorial stress combination on plants, crops, and ecosystems: how should we prepare for what comes next? *The Plant Journal, 117*(6), 1800–1814. https://doi.org/10.1111/tpj.16557

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., & Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences, 114*(35), 9326–9331. https://doi.org/10.1073/pnas.1701762114

Zheng, G., Chen, J., & Li, W. (2020). Impacts of CO_2 elevation on the physiology and seed quality of soybean. *Plant Diversity, 42*(1), 44–51. https://doi.org/10.1016/j.pld.2019.09.004

Zinta, G., AbdElgawad, H., Peshev, D., Weedon, J. T., Van den Ende, W., Nijs, I., … Asard, H. (2018). Dynamics of metabolic responses to periods of combined heat and drought in Arabidopsis thaliana under ambient and elevated atmospheric CO_2. *Journal of Experimental Botany, 69*(8), 2159–2170. https://doi.org/10.1093/jxb/ery055

Zipper, S. C., Qiu, J., & Kucharik, C. J. (2016). Drought effects on US maize and soybean production: Spatiotemporal patterns and historical changes. *Environmental Research Letters, 11*(9), Article 094021. https://doi.org/10.1088/1748-9326/11/9/09402