Developing an eco-physiological process-based model of soybean growth and yield (MATCRO-Soy v.1): Model calibration and evaluation

Astrid Yusara[1,8], Tomomichi Kato[2]*, Elizabeth A. Ainsworth[3]*, Rafael Battisti[4], Etsushi Kumagai[5], Satoshi Nakano[6], Yushan Wu[7], Yutaka Tsusumi-Morita[8], Kazuhiko Kobayashi[8], Yuji Masutomi[8]

[1]Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan
[2]Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan
[3]Departments of Crop Sciences and Plant Biology, University of Illinois Urbana-Champaign, Urbana, Illinois, USA
[4]Escola de Agronomia, Universidade Federal de Goiás, Goiânia, Brazil
[5]Division of Agro-Environment Research Agricultural Meteorology Group, Tohoku Agricultural Research Center, NARO, Tsukuba, Ibaraki, Japan
[6]Central Region Agricultural Research Center, NARO, Tsukuba, Ibaraki, Japan
[7]College of Agronomy, Sichuan Agricultural University, Chengdu, PR China
[8]Center for Climate Change Adaptation, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan

Correspondence to: Tomomichi Kato (tkato@agr.hokudai.ac.jp)

Deleted: 060-0808
Deleted: 060-0808
Deleted: 305-8666
Deleted: 611130,
Deleted: of

Abstract. MATCRO-Soy is an eco-physiological process-based crop model for soybean (*Glycine max* L. (Merr.)). It was developed by modifying the parameters of MATCRO-Rice, integrates crop growth processes with a land surface model. These modifications were made using data from literature and field experiments across the world. The reliability of the model was validated extensively by observed soybean yield data across the global, national, and grid cell levels. A moderate correlation was observed between the MATCRO-Soy and FAOSTAT yield data with correlation coefficients of 0.81 ($p < 0.001$) for the global average yield and 0.45 ($p < 0.01$) for the global average detrended yield over a 34-year period (1981-2014). Furthermore, the grid-cell level validation revealed that 66 % of the grid cells in the global yield map exhibited a statistically significant correlation between the MATCRO-Soy simulated yield and the reference data derived from observational records. These results highlight the model's ability to reproduce soybean yield under different environmental conditions, integrating soil water availability and nitrogen fertilizer. MATCRO-soy could enhance our understanding of crop physiology, especially crop responses to climate change. Its application may support efforts to reduce uncertainty in projections of climate change impacts on soybeans.

Deleted: . The original model, MATCRO-Rice
Deleted: 512
Deleted: 71
Deleted: ,
Deleted: response
Deleted: and

1 Introduction

Crop growth models have been widely used for yield estimation, agricultural management practice optimization, climate change impact evaluation, and informing decision-making about food security strategies (Adeboye et al., 2021; Cuddington et al., 2013; Hoogenboom, 2000). Given the significant impact of weather variability on global yield (Müller et al., 2017; Ray et al., 2015), process-based models can represent the long-term climate change impacts on productivity via the influence of key climatic factors on physiological processes that are represented in the model (Boote et al., 2013; Cuddington et al., 2013; Fodor et al., 2017; Jones et al., 2017; Marin et al., 2014; Stöckle and Kemanian, 2020). Process-based models explicitly incorporate the crucial eco-physiological processes of photosynthesis and stomatal conductance, improving predictions under varying climate scenarios compared with models that focus on the empirical relationship between absorbed radiation and assimilation through radiation use efficiency (Jin et al., 2018). Hence, crop models are useful for capturing the complexity of

Deleted: changes
Deleted: physiological processes influenced by key climate factors on …
Deleted: of climate
Deleted: yield
Formatted: Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Formatted: Font color: Black
Deleted: mechanistic crop
Deleted: direct

soil-crop-climate interactions for ensuring food security, optimizing yields, promoting sustainability, and planning adaptation strategies (García-Tejero et al., 2011). Global-scale simulations are essential to enhance these efforts by understanding interactions between physiological processes and environmental factors, supporting adaptive management practices and strengthening agricultural resilience.

The Agricultural Model Intercomparison and Improvement Project (AgMIP) has examined the performance of global gridded crop models (GGCMs) in simulating the potential impact of climate change on crop yield (Müller et al., 2017; Kothari et al., 2022). AgMIP has demonstrated that the simulated impacts of environmental factors on crop yields using a GGCM generally align with measurements and that a model ensemble reduces uncertainty (Elliott et al., 2015). However, yield change under future climate change scenarios shows inconsistent results and greater variability in soybean than in other crops because of model discrepancies (Jägermeyr et al. 2021). Despite being a major crop, soybean (*Glycine max* L. (Merr.)), has been studied less extensively than other crops in terms of crop response to changing environments (Ruane et al., 2017; Kothari et al., 2022). Therefore, the development of a new soybean model is needed to reduce uncertainties in climate change impact assessments.

It is important to utilize a diverse type of crop models and ensure model diversity to accurately understand the uncertainties of simulations, as relying on a single model can lead to biased results. To our knowledge, only five process-based models for global-scale soybean yield estimation with leaf-level photosynthesis and stomatal conductance parameters exist, including LPJ-GUESS (Ma et al., 2022), LPJmL (Wirth et al., 2024), ORCHIDEE-crop (Wu et al., 2016), PRYSBI2 (Sakurai et al., 2014), and JULES (Leung et al., 2020), making this approach relatively uncommon. Thus, further development and validation of process-based models that incorporate leaf-level photosynthesis and stomatal conductance parameters are essential.

MATCRO (Masutomi et al., 2016a), is an ecosystem process-based model for crops embedded into the land surface model of minimal advanced treatments of surface interaction and runoff (MATSIRO; Takata et al., 2003) with a crop growth model for rice, which is further explained in Section 2. MATCRO-Rice uses state variables to exchange information (e.g. temperature, soil moisture, transpiration, leaf area index, and photosynthesis rate) between the land surface model and crop growth model. The mechanisms that consider photosynthesis and stomatal conductance to assess the impact of greenhouse gases on carbon and water fluxes have been incorporated into MATCRO-Rice. Masutomi et al. (2019) described the implementation of ozone effects within these mechanisms, indicating the model's capability to account for environmental stressor. Furthermore, MATCRO-Rice has been applied at the regional scale, and it has been utilized to measure climate impacts, which are important for developing adaptation strategies (Kinose and Masutomi, 2020; Masutomi, et al., 2016b).

We developed a new process-based model for soybean (MATCRO-Soy v.1) that incorporates diverse biological processes and environmental interactions that drive plant growth and adaptation to changing conditions. Adapted from MATCRO Rice, this model is applied for soybeans by parameterizing key processes using experimental data and findings from the literature. The current version of MATCRO-Soy (v.1) was evaluated in a global-scale simulation, following a calibration process that considered essential photosynthesis mechanisms. This paper presents the model description in Section 2, the calibration process in Section 3, and the model evaluation in Sections 4 and 5.

2 Model Description

MATCRO-Soy is based on MATCRO-Rice, a process-based model of rice growth and yield, which has been modified for use in soybeans. MATCRO-Rice is a combined land surface and crop growth model used to explore land-atmosphere interaction in rice fields. Unlike the MATCRO-Rice v.1 version, MATCRO-Soy focuses on yield simulation and omits the calculation of sensible and latent heat fluxes in the energy balance to reduce computational complexity while maintaining accuracy in simulating soybean growth and yield.

Deleted:

Formatted: Font color: Black

Deleted: efforts have

Deleted: estimated

Deleted: using a GGCM

Deleted: due

Deleted:

Deleted: the

Deleted: crucial for reducing

Deleted: remains

Deleted:

Deleted: Crop growth

Deleted: , which are widely used

Deleted: (e.g. ozone) in Masutomi et al. (2019),

Deleted: initially

Deleted: only

2.1 Overview of MATCRO-Soy

MATCRO-Soy includes three main modules: phenology, photosynthesis, and carbon partitioning (Figure 1). The phenology module simulates crop phenological development over time based on heat unit accumulation. The module directs the progression of carbon assimilation and partitioning by monitoring plant developmental stages from sowing to harvest. The phenology module simulates developmental stages based on developmental rate from sowing to harvest and influences key processes such as glucose production and allocation across plant organs. The photosynthesis module initially estimates gross primary production (GPP) and respiration at the leaf-level using the Farquhar model (Farquhar et al., 1980), and extends net primary production (NPP) estimation to the canopy level following the concept of de Pury and Farquhar (1997). It considers the electron-transport-limited rate of photosynthesis, Rubisco-limited photosynthesis, and leaf respiration to estimate NPP at leaf-level.

The photosynthesis and carbon partitioning modules are closely linked, as carbon assimilation from photosynthesis is subsequently allocated to different plant organs. The NPP is stored in glucose and starch reserves. The carbon partitioning module distributes the glucose into each organ (i.e. leaf, stem, root, and storage organ) following the method derived from the school of de Wit by simulating biosynthetic processes (de Vries et al., 1989). MATCRO accounts for leaf senescence as it influences nutrient cycling, crop productivity, and the leaf area index, which plays an important role in canopy photosynthesis. Leaf senescence is simulated as a function of crop developmental stage, as defined by the phenology module. MATCRO incorporates the amount of nitrogen per leaf area (specific leaf nitrogen) as a key determinant of photosynthetic capacity. Root depth can affect photosynthesis indirectly through the plant's ability to access water and nutrients from soil layers, further influencing plant growth within the model framework.

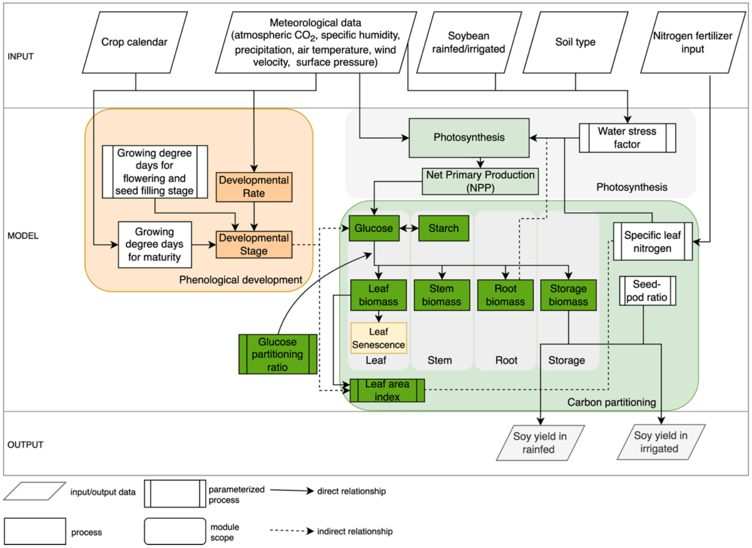


Figure 1. Flowchart diagram of soybean yield simulation by MATCRO-Soy.

Deleted: Figure 1

Deleted: The photosynthesis and carbon partitioning modules are closely linked with the carbon allocation driven by photosynthetic activity. The

Formatted: Font: Not Bold

Deleted: serves as a

Deleted: dimension

Deleted: and

Deleted: of the processes

Deleted: includes the absorbed photosynthetically active radiation (PAR) in

Deleted:

Deleted: to produce the net primary product (NPP). These

Formatted: Font color: Black

Deleted: products

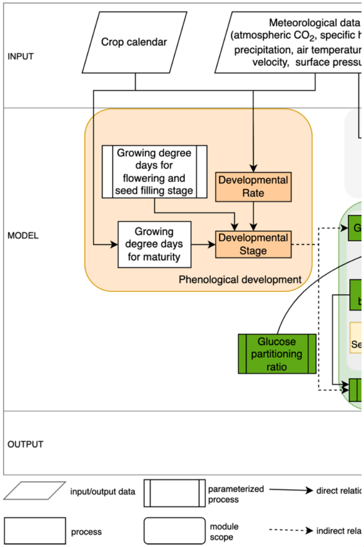
Deleted: proposed by

Deleted: De

Deleted: used in MACROS

Deleted: also driven

Deleted:



The input data consisted of environmental variables obtained from meteorological forcings, soil type classifications, nitrogen fertilizer applications, and agricultural management practices such as irrigation and seed sowing. These inputs are crucial for setting the initial conditions and boundary parameters for the simulations. The output of MATCRO is the crop yield (kg ha^{-1}) estimated for both irrigated and rainfed conditions on the basis of soil-crop interactions. First, we processed the parameterized growing degree days for maturity using crop calendar data to estimate the harvest time in the phenology module (see section 2.2). The photosynthesis module includes limiting factors such as nitrogen fertilization and water stress, as detailed in Section 2.3. Then, the crop growth is calculated based on its developmental stage (Section 2.4). We conducted a parameterization process encompassing phenological development, carbon partitioning, and photosynthesis limited by water stress and nitrogen uptake. The crop yield was estimated using the parameterized seed-pod ratio (see section 2.5). The adjusted parameters in MATCRO-Soy are described in Section 2.6 where the key dynamic variables are parameterized over time to ensure reliable carbon assimilation in soybean. This comprehensive approach allows MATCRO to account for complex interactions between environmental conditions, crop physiology, and management practices, providing a robust framework for predicting crop yield and assessing agricultural productivity.

2.2 Crop phenological development

Phenological development defines the timing of developmental events based on environmental inputs. MATCRO calculates crop developmental stages (DVS) using an index indicating the sowing time (DVS=0) to maturation time (DVS=1) on the basis of the integral of the temperature required to exceed the phenological changes. The module uses a formulation based on Bouman et al. (2001) as outlined in Equations (1) to (4).

$$DVS_t = GDD_t / GDD_m \tag{1}$$

$$GDD_t = \int_0^t DVR dt' \tag{2}$$

$$GDD_m = \int_0^m DVR dt' \tag{3}$$

$$DVR_t = \begin{cases} 0, & T_t < T_b \mid T_t > T_h \\ T_t - T_b, & T_b < T_t < T_o \\ \frac{(T_o - T_b)(T_h - T_t)}{(T_h - T_o)}, & T_o < T_t < T_h \end{cases} \tag{4}$$

where GDD_t and GDD_m indicate the growing degree days (°C days) used to estimate the development of plants during the growing season at time t and at maturity, respectively. DVR represents the developmental rate at t, whereas T_t represents the temperature at t. The parameters T_b, T_o, and T_h (°C) are crop-specific and represent the base, optimum, and highest temperatures for crop development, respectively.

The impact of temperature on phenological stages varies for each crop stage as Boote et al. (1998) observed that cardinal temperatures (T_b, T_h, T_o) may differ for vegetative and reproductive stages. We follow de Vries et al. (1989) during the growing season to simplify the calculation input and also because more detailed data in each phenological stage is lacking. This study parameterized the developmental stages at flowering (DVS_f), seed filling (DVS_s), and maturation (DVS_m) stages on the basis of the experimental datasets by calculating the mean, values listed in Table 2. MATCRO uses these DVS parameters to define the period of leaf dry weight loss due to leaf senescence and the remobilization of starch reserves from the stem (Masutomi et al. 2016a). We assume that this phenological time in soybean is in the middle of the flowering and seed filling stage parameterized in this study as leaf loss started within those periods.

2.3 Carbon assimilation process

In the photosynthesis module of MATCRO-Soy, carbon assimilation is based on leaf-level photosynthesis calculations in sunlit and shaded conditions (Dai et al., 2004) to predict canopy photosynthesis. The calculation includes the stomatal conductance

Deleted: the

Deleted: module was calibrated

Deleted: $DVS = GDD/GDD_m$

Deleted: $GDD = \int_0^t DVRdt'$

Deleted: DVR

Deleted: $\begin{cases} 0, & T_m < T_b \mid T_m > T_h \\ T_m - T_b, & T_b < T_m < T_o \\ \frac{(T_o - T_b)(T_h - T_i)}{(T_h - T_o)}, & T_o < T_m < T_h \end{cases}$

Deleted: GDD

Deleted: T_m

Deleted: T_o

Deleted: minimum

Deleted: maximum

Deleted: due to the lack of

Deleted: the

Deleted: Table 2.

Deleted: Table 2

Formatted: Font: 10 pt, Not Bold

Deleted: the

response to relative humidity (Collatz et al., 1991). The net carbon assimilation (A_n) in MATCRO is calculated using the Farquhar model as further described in Masutomi et al. (2016a), expressed in Eq. (5).

$$A_n = f(PAR, \mathrm{P_a}, T_{leaf}, CO_{2leaf}, V_{\mathrm{cmax}}, BB_a, BB_b) \tag{5}$$

A_n (mol(CO$_2$) m^{-2} s^{-1}) represents net carbon assimilation that contributes to NPP for biomass growth. It is a function of the intensity of absorbed photosynthetic active radiation (PAR, in mol(photon) m^{-2} s^{-1}), air pressure ($\mathrm{P_a}$, in Pa), leaf temperature (T_{leaf}, in K), CO$_2$ concentration at the substomatal chamber (CO_{2leaf}, in Pa(CO$_2$) Pa(Air)$^{-1}$), maximum Rubisco activity (V_{cmax}, in mol(CO$_2$) m^{-2} s^{-1}), the slope (BB_a, in mol(H$_2$O)m^{-2}s^{-1}) and intercept (BB_b, in mol(H$_2$O) m^{-2} s^{-1}) of the Ball-Berry model of the relationship between crop assimilation, stomatal conductance per unit leaf area, relative humidity at the leaf surface, and ambient CO$_2$ concentration (Ball, 1988). In this study, we assume the leaf temperature is the same as air temperature to reduce the complexity of the calculation.

Rubisco activity (V_{cmax}) is a key variable used to assess the carbon rate entering the photosynthetic pathway, as it catalyzes the crucial initial step of RuBP (Ribulose-1,5-bisphosphate) carboxylation in photosynthetic carbon assimilation for C3 plants (Sage, 2002; Xu et al., 2022). In MATCRO, V_{cmax} is calculated as follows:

$$V_{\mathrm{cmax}} = V_{\mathrm{ctop}} \exp\left(-K_{\mathrm{n}} LAI\right) \tag{6}$$

$$V_{\mathrm{ctop}} = \max\left(aSLN^2 + bSLN + c, V_{ctopmax}\right) \tag{7}$$

V_{cmax} is the maximum Rubisco activity within the canopy (mol(CO$_2$) m^{-2} s^{-1}) limited by the exponential value of vertical distribution of leaf nitrogen (K_{n}), leaf area index (LAI, in m^2 m^{-2}), and maximum Rubisco activity at the top of canopy (V_{ctop}, in mol(CO$_2$) m^{-2} s^{-1}). We determined the V_{ctop} for photosynthetic rate limited by the specific leaf nitrogen (SLN) in Eq. (7) for soybean using the relationship between two parameters of rubisco activity and leaf nitrogen from experiments summarized from Ainsworth et al. (2014) in the reproductive stage and Qiang et al. (2022) in the vegetative stage. This relationship is empirically represented with a polynomial quadratic equation limited by maximum value of Rubisco activity at the top canopy (V_{ctop} in mol(CO$_2$) m^{-2} s^{-1}). a, b, c are quadratic coefficient, linear coefficient, and constant respectively from the relationship of both variables where the data has been digitized from WebPlotDigitizer (Rohatgi, 2023).

MATCRO considers nitrogen fertilization input denoted as N_{fert} (unit: kg(N) ha^{-1}) which influences the amount of specific leaf nitrogen (SLN, g(N) m^{-2}), particularly under conditions of limited nitrogen availability (La Menza et al., 2023; Thies et al., 1995). SLN was determined by nitrogen supply (including biological nitrogen fixation, soil mineral nitrogen, and nitrogen fertilizer) and by plant demand. The changes in SLN over the growing period in MATCRO-Soy simulated a function derived from La Menza et al. (2023) which observed SLN under wide range of low and high nitrogen fertilization conditions (see Supplementary file Figure S1). It shows the specific leaf nitrogen value toward the crop growth period where higher nitrogen fertilizer results in a higher leaf nitrogen content. In the absence of empirical data for initial growth stages, the model assumes a gradual increase in nitrogen content. The simulated SLN under different nitrogen fertilization treatments is described in Eq. (8) and (9).

$$SLN = \begin{cases} SLN_{Y0} + \frac{(SLN_{Y1}-SLN_{Y0})(DVS-SLN_{X1})}{SLN_{X1}}, \; if\ DVS < SLN_{X1} \\ SLN_{Y2} + \frac{(SLN_{Y2}-SLN_{Y1})(DVS-DVS_f)}{(DVS_f-SLN_{X1})}, \; if\ SLN_{X1} \le DVS < DVS_f \\ Y + \frac{(Y-SLN_{Y2})(DVS-DVS_s)}{(\text{▯▯}VS_s-DVS_f)}, \; if\ DVS_f \le DVS < DVS_s \\ SLN_{Y0} + \frac{(SLN_{Y0}-Y)(DVS-DVS_m)}{(DVS_m-DVS_s)}, \; if\ DVS_s \le DVS \le DVS_m \end{cases} \tag{8}$$

$$\mathrm{Y} = SLN_{Y3,l} + \frac{SLN_{Y3,h}-SLN_{Y3,l}}{N_{fert,high}} * N_{fert} \tag{9}$$

SLN values vary across different phenological stages, with the developmental stage (DVS) value of soybean growth ranges from 0 (sow) to 1 (harvest). We define DVS_f, DVS_s, DVS_m, and SLN_{X1} as the start of flowering, seed filling, maturity time, and

Deleted: net primary product

Deleted: V_{cmax}

Deleted: i

Deleted: at the top of

Deleted: (µ

Deleted: and

Deleted: ,

Deleted: µ

Deleted: (

Deleted: ;

Deleted: .

Deleted: change

Deleted: follows the study of (

Deleted: .,

Deleted: measured

Deleted: , as

Deleted: $\frac{SLN_{Y3,h}-SLN_{Y3,l}}{N_{fert,max}} * N_{fert}$

the point where the SLN pattern started to changes with the parameterized values of 0.4, 0.659, 1, and 0.15 respectively. While SLN_{Y0}, SLN_{Y1}, SLN_{Y2}, $SLN_{Y3,h}$, $SLN_{Y3,l}$ represent the SLN at the time of initial stage, early decline, pre-flowering increase, subsequent decline during the reproductive stage under no and high (h) nitrogen inputs with the value of 0.75, 2.25, 1.7, 0.75, and 1.8, respectively. $N_{fert,high}$ refers to the high nitrogen fertilizer input used in the model for parameterization, as described in Table 2. Y denotes the observed gap function in specific leaf nitrogen under high and low nitrogen fertilizer treatments (g(N) m^{-2}) in Supplementary file Figure S1.

The growth stages are parameterized based on experimental datasets and align with study from Irmak et al. (2013) using the growth stage classification by Fehr and Caviness (1977). SLN primarily depends on nitrogen derived from biological fixation and soil nitrogen, either from natural sources or applied fertilizers. Nitrogen uptake, including biological nitrogen fixation and uptake from soil nitrogen, is implicitly captured through SLN that influence V_{cmax} in Eq. (7) and (8), while the effect of applied fertilizers in Eq. (8) and (9).

2.4 Crop growth dynamics

The products of photosynthesis contribute to glucose reserves, which provide energy for growth during various developmental stages. The crop growth dynamics include a carbon biomass partitioning module to calculate the dry weight of each soybean organ (W_{organ} in kgha^{-1}). This variable is the accumulated value of growth rate of dry weight (G_{organ} in kg ha^{-1} s^{-1}) during the time from emergence to harvest. Further details on this module can be found in Masutomi et al. (2016a).

$$W_{organ} = f(G_{organ}) \tag{10}$$

We calculate the W_{organ} in each soybean organ (i.e. leaf, stem, pod including the seed, glucose reserves and starch). Growth rate of the dry weight (G_{organ} in kg ha^{-1} s^{-1}) is calculated based on the parameters of conversion factor of dry weight from glucose to organ ($F_{glu-organ}$ in kgha^{-1}(kg ha^{-1})$^{-1}$) for leaf, stem, pod, root, and starch (listed in Table 1), and ratio of glucose partitioned to organ (P_{organ}) for shoot, leaf, and pod (listed in Table 2). Shoot refers to aboveground biomass parts including the stem, leaf, and pod. G_{organ} for each organ and storage, leaf, pod, root, stem, and starch, are expressed below:

$$G_{glu} = f(W_{leaf}, A_{glu}, R_{glu}) \tag{11}$$

$$G_{leaf} = G_{glu} P_{shoot} P_{leaf} F_{glu\text{-}leaf} \tag{12}$$

$$G_{stem} = G_{glu} P_{shoot} (P_{leaf} - P_{pod}) \times (1 - f_{starch}) F_{glu-stem} \tag{13}$$

$$G_{pod} = G_{glu} P_{shoot} P_{pod} F_{glu-pod} \tag{14}$$

$$G_{root} = G_{glu} (1 - P_{shoot}) F_{glu-root} \tag{15}$$

$$G_{starch} = G_{glu} P_{shoot} (P_{leaf} - P_{pod}) f_{starch} F_{glu-starch} \tag{16}$$

G_{glu} (kg ha^{-1} s^{-1}) is the amount of glucose partitioned to soybean organ and reserve derived from function of dry weight of leaf (W_{leaf} in kg ha^{-1}), net carbon assimilation in glucose form (A_{glu} in kg(CH$_2$O) ha^{-1} s^{-1}), and the remobilization from starch reserve in the stem after conversion to glucose (R_{glu} in kg ha^{-1} s^{-1}). A_{glu} is A_n that has been already converted using the conversion factor from CO$_2$ to glucose using the value of 1.08×10^6 [kg ha^{-1} h^{-1}(mol m^{-2} s^{-2})$^{-1}$] that is the physical and chemical constant for the conversion. R_{glu} is the remobilization from starch reserve in the stem after converted to glucose using ratio of remobilization value. This R_{glu} is subtracted from the dry weight of starch reserves (W_{starch}). f_{starch} [kg ha^{-1}(kg ha^{-1})$^{-1}$] is the fraction of glucose allocated to starch reserves calculated in stem dry weight loss. Each growth rate of dry weight (G_{organ}) is calculated based on the parameters conversion factor of dry weight ($F_{glu-organ}$) and ratio of glucose partitioned to organ (P_{organ}) value as follow in Eq(17) –(19):

Deleted: These

Formatted: Font: 10 pt, Not Bold

Deleted: Table 2

Formatted: Indent: First line: 0.5 cm

Formatted: Font color: Auto

Formatted: Normal, Right: 0 cm

Formatted: Font color: Auto

Formatted: Font: Times New Roman, Font color: Auto

Formatted: Font color: Auto

Deleted:)

Formatted: Font color: Auto

Formatted: Font: Times New Roman, Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Times New Roman, Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Times New Roman, Font color: Auto

Deleted: Table 1

Formatted: Font color: Auto

Deleted: in kg ha^{-1}s^{-1})

Deleted: Table 2

Deleted: Table 2

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Not Bold

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Times New Roman, Font color: Auto

Formatted: Font: Times New Roman, Font color: Auto

Deleted: 1.08x10^6

Deleted:

Deleted:

$$P_{shoot} = \begin{cases} 1 - P_{root}, \ \ if\ DVS = 0 \\ \frac{1 - P_{root}(DVS_m - DVS)}{DVS_m}, \ if\ 0 < DVS < DVS_m \\ 1, \ \ if\ DVS \geq DVS_m \end{cases} \tag{17}$$

$$P_{leaf} = \begin{cases} P_{leaf0} + \frac{DVS}{DVS_{leaf1}}(P_{leaf1} - P_{leaf0}), \ \ if\ DVS < DVS_{leaf1} \\ P_{leaf2} - \frac{(P_{leaf2} - P_{leaf1})}{DVS_{leaf2} - DVS_{leaf1}}(DVS_{leaf2} - DVS), \ if\ DVS_{leaf1} \leq DVS < DVS_{leaf2} \\ 0, \ \ if\ DVS \geq DVS_{leaf2} \end{cases} \tag{18}$$

$$P_{pod} = \begin{cases} 0, \ \ if\ DVS < DVS_{pod1} \\ \frac{DVS - DVS_{pod1}}{DVS_{pod2} - DVS_{pod1}}, \ if\ DVS_{pod1} \leq DVS < DVS_{pod2} \\ 1, \ \ if\ DVS \geq DVS_{pod2} \end{cases} \tag{19}$$

P_{leaf0}, P_{leaf1}, P_{leaf2} represent the glucose partitioning ratio of leaf toward shoot at the time of initial stage when leaf growth starts to decline (leaf0), stop growing (leaf1), and at maturity (leaf2), respectively. While DVS_{pod1} and DVS_{pod2} indicate the developmental stage values at which glucose partitioning of pod to the shoot begins to increase and eventually saturates (Figure 2). Figure 2 in section 3.2 visually represented the glucose partitioning ratio during crop growth as calibrated in this study.

The glucose partitioned in each organ is adjusted during the developmental stage using experimental data in the calibration process, further described in Section 3. However, the dry weight of leaf in this module is reduced due to leaf senescence by calculating loss of leaf dry weight (L_{leaf} in kg ha^{-1} s^{-1}) derived from the calibration of partitioned glucose ratio to the ratio of dead leaf (P_{dleaf} in s^{-1}), as outlined in Eq. (20) and (21).

$$L_{leaf} = \begin{cases} 0 \qquad\qquad , \ if\ DVS < DVS_{deadleaf1} \\ P_{dleaf}(W_{leaf} - W_{glu}), \ if\ DVS \geq DVS_{deadleaf1} \end{cases} \tag{20}$$

$$P_{dleaf} = P_{deadleaf2} \frac{(DVS - DVS_{deadleaf1})}{(1 - DVS_{dealeaf1})} \tag{21}$$

Then we calculate the leaf area index (LAI) that serves as a parameter to assess the leaf surface area relative to the ground area. It directly influences the plant ability to intercept solar radiation for photosynthesis.

LAI is computed as follow:

$$LAI = \frac{W_{leaf} + W_{glu}}{SLW} \tag{22}$$

LAI is calculated from the estimated leaf dry weight (W_{leaf}, in kg ha^{-1}) and glucose reserves in leaves (W_{glu}, in kg ha^{-1}) divided with specific leaf weight (SLW, in kg ha^{-1}). SLW indicates leaf dry weight per unit leaf area. The value of SLW dynamically changed during the developmental stage following exponential relationship:

$$SLW = SLW_{max} + (SLW_{min} - SLW_{max})\ exp(-SLW_x DVS) \tag{23}$$

SLW_{max}, SLW_{min}, and SLW_x represent the maximum, minimum, and slope parameters, respectively, that define the values observed in the exponential relationship based on experimental dataset in Table 3. In addition to LAI, photosynthesis is also indirectly affected by the root depth (z_{root}, in m) that determines the plant capacity for water and nutrient uptake. Root depth is calculated as follow:

$$z_{root} = f(r_{root}, z_{rootmax}) \tag{24}$$

z_{root} is the accumulative value from growth rate of root depth (r_{root}, in mm day^{-1}) limited by maximum possible root depth ($z_{rootmax}$, in meter).

2.5 Soybean yield estimation

The soybean yield is calculated from the pod dry weight at harvest ($W_{podharvest}$, in kg ha^{-1}) via the seed-pod ratio (SR) in MATCRO-Soy. The yield is further affected by water stress ($f_{wstress}$) in Eq. (25).

Deleted: P_{shoot} alternatives: $1 - P_{root}, \ if\ DVS < 0$; $\frac{1 - P_{root}(DVS_m - DVS)}{DVS_m}, \ if\ 0 \leq DVS < DVS_m$; $1, \ if\ DVS \geq DVS_m$

Deleted: using the adjusted specific leaf weight (SLW in kg ha^{-2}) expressed

Deleted: ^{2}

Deleted: weight (

Deleted: ^{2}

Deleted: parameterized

Deleted: a key parameter that links

Deleted: to surface

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Deleted: Table 3.

Deleted: Table 3

Formatted: Font: 10 pt, Not Bold

Formatted: Font: Times New Roman

$$Yield = f(W_{podharvest}, SR, f_{wstress}, T) \tag{25}$$

The yield was calculated using the parameter *SR*, which is the ratio of yield (seed, kg ha^{-1}) to the storage organ of the pod ($W_{podharvest}$) at harvest time and was derived from experimental datasets in Table 3. *T* is the temperature (Kelvin) that limits heat and cold damage to the yield of soybean. The water stress factor ($f_{wstress}$) was determined on the basis of the fraction of available soil water at the soil layer -i (FAW_i) over crop yield in timestep *t* during the crop growth, based on a previous study on the relationship between the soybean transpiration ratio and transpirable soil water conducted by Ray and Sinclair (1998), given in Eq (26).

$$f_{wstress,t} = \begin{cases} \frac{1}{0.5} FAW_i, & if\ FAW_i \leq 0.5 \\ 1, & if\ FAW_i > 0.5 \end{cases} \tag{26}$$

The value of $f_{wstress,t}$ depends on soil water availability at soil layer-i (FAW_i), which is the estimated soil water content based on the water flux between the soil layers (Masutomi et al., 2016a) during the crop growth calculated via Eq. (27):

$$FAW_i = \frac{WSL_i - WSL_{wilt}}{WSL_{FC} - WSL_{wilt}} \tag{27}$$

where WSL_i, WSL_{wilt}, and WSL_{FC} represent the water level in the soil layer -i, wilting point, and field capacity, respectively. A value of $f_{wstress}$ equal to 1 indicates no water stress as the fraction of available soil water is adequate for crop growth. Hence, yield is calculated as the potential yield constrained by water stress.

2.6 Soybean-specific parameters

MATCRO-Soy shares several parameters with MATCRO-Rice as both are C3 species. However, soybean differs from cereal crops because of its nitrogen-fixing ability. This characteristic is represented through specific leaf nitrogen during the crop growth, as described in Eqs. (8) and (9). The crop-specific parameters reflect the unique physiological and chemical processes involved in soybean growth. but still align with the general framework of MATCRO-Rice. Key parameter adjustments are outlined in Table 1 as MATCRO employs a set of specific parameters to simulate crop growth and yield. These parameters include factors related to carbon allocation, root growth characteristics, and crop development based on cardinal temperatures. By accurately representing the unique physiological and biochemical characteristics of soybeans, these parameters contribute to the ability of the model to predict crop yield with greater precision.

MATCRO-Soy aims for simulations applicable to a global scale; hence, it uses a single global parameterization as a standardized set of parameters applied worldwide. It uses a unified approach for modelling crop behaviour across different regions. It was assumed that the parameter values from the different treatments and cultivars were independent. Table 2 contains a list of variables parameterized within the model, including the glucose partitioning, nitrogen parameters, and photosynthetic capacity. Through the parameterization of these variables, the model can be adapted for various growing conditions and employed to assess the sensitivity of crop performance to different factors. These parameters are commonly used to evaluate the crop model sensitivity to environmental changes and require further attention, as highlighted by simulations from other crop model as wells (Battisti et al., 2018a).

Table 1. Crop-specific parameters used for MATCRO-Soy

Parameters	Description	Value	Units	Source	Eq.
$F_{glu-leaf}$	conversion factor of dry weight from glucose to leaf	0.871	kg ha^{-1} (kg ha^{-1})$^{-1}$	de Vries et al. (1989)	(12)
$F_{glu-stem}$	conversion factor of dry weight from glucose to stem	0.810	kg ha^{-1} (kg ha^{-1})$^{-1}$	de Vries et al. (1989)	(13)
$F_{glu-root}$	conversion factor of dry weight from glucose to root	0.857	kg ha^{-1} (kg ha^{-1})$^{-1}$	de Vries et al. (1989)	(15)
$F_{glu-pod}$	conversion factor of dry weight from glucose to pod	0.759	kg ha^{-1} (kg ha^{-1})$^{-1}$	de Vries et al. (1989)	(14)

Deleted: , kg ha^{-1})

Deleted: Table 3

Deleted: Table 3

Formatted: Font: Not Bold

Deleted: $f_{wstress}$

Deleted: $f_{wstress}$

Deleted: **Table 1**

Deleted: Table 1

Formatted: Font: Not Bold

Deleted: Table 2

Deleted: Table 2

Formatted: Font: Not Bold

Deleted: 2018)

Deleted: Penning

Deleted: Penning

Deleted: Penning

Deleted: Penning

Parameters	Description	Value	Units	Source	Eq.
$F_{glu-starch}$	carbon fraction in the dry matter of starch	0.9	kg ha^{-1} (kg ha^{-1})$^{-1}$	Physical and chemical constant	(15)
K_N	vertical distribution of leaf nitrogen	0.11	-	Bonan et al. (2011)	(6)
r_{root}	rate of root depth increase	0.03	mm day^{-1}	Ordóñez et al. (2018) ; Nakano et al. (2021)	(24)
$Z_{rootmax}$	maximum root depth	1.7	m	de Vries et al. (1989)	(24)
T_b	base temperature for crop development	10	℃	de Vries et al. (1989)	(4)
T_h	highest temperature for crop development	34	℃	de Vries et al. (1989)	(4)
T_o	optimum temperature for crop development	27	℃	de Vries et al. (1989)	(4)

Deleted: Penning

Deleted: Penning

Deleted: Penning

Deleted: Penning

Formatted: Font: Bold

Formatted: Caption, Justified, Indent: First line: 0 cm

Table 2. Parameterized variables for soybean in MATCRO

Variables	Value	Units	Description
a	-18.516	-	coefficient at relationship of rubisco activity and leaf nitrogen in Eq. (7)
b	114.33	-	coefficient at relationship of rubisco activity and leaf nitrogen in Eq. (7)
c	-73.336	-	constant at relationship of rubisco activity and leaf nitrogen in Eq. (7)
$DVS_{deadleaf1}$	0.6	-	1st DVS point where the dead leaf ratio pattern changes
$DVS_{deadleaf2}$	1	-	2nd DVS point where the dead leaf ratio pattern changes
DVS_f	0.4	-	developmental stage on initial flowering stage
DVS_{leaf1}	0.25	-	1st DVS point where the leaf partitioning pattern changes
DVS_{leaf2}	0.659	-	2nd DVS point where the leaf partitioning pattern changes
DVS_m	1	-	developmental stage at maturity time
DVS_{pod1}	0.48	-	1st DVS point where the pod partitioning pattern changes
DVS_{pod2}	0.72	-	2nd DVS point where the pod partitioning pattern changes
DVS_s	0.659	-	developmental stage to start seed filling stage
DVS_{SLN1}	0.4	-	1st DVS point where the specific leaf nitrogen changes along with DVS
DVS_{SLN2}	0.4	-	2nd DVS point where the specific leaf nitrogen changes along with DVS
DVS_{SLN3}	0.659	-	3rd DVS point where the specific leaf nitrogen changes along with DVS
f_{starch}	0.18	-	fraction of glucose allocated to starch reserves
SR	0.68	-	seed-pod ratio (SR) accounting harvest index from storage organ
$N_{fert,high}$	300	$kgNha^{-1}$	nitrogen fertilizer value used in high nitrogen fertilizer in La Menza et al. (2023)
P_{leaf0}	0.38	-	glucose partitioning ratio of leaf toward shoot in the initial DVS point
P_{leaf1}	0.6	-	glucose partitioning ratio of leaf toward shoot in the 1st DVS point
P_{leaf2}	0	-	glucose partitioning ratio of leaf toward shoot in the 2nd DVS point
$P_{deadleaf1}$	0	s^{-1}	dead leaf ratio value in the 1st DVS point
$P_{deadleaf2}$	0.000001	s^{-1}	dead leaf ratio value the 2nd DVS point

Deleted: 53

Variables	Value	Units	Description
SLN_{Y0}	0.75	gNm^{-2}	initial specific leaf nitrogen
SLN_{Y1}	2.25	gNm^{-2}	specific leaf nitrogen value in the 1st DVS point
SLN_{Y2}	1.7	gNm^{-2}	specific leaf nitrogen value in the 2nd DVS point
$SLN_{Y3,h}$	0.75	gNm^{-2}	specific leaf nitrogen value in the 3rd DVS point when using high nitrogen fertilizer
$SLN_{Y3,l}$	1.8	gNm^{-2}	specific leaf nitrogen value in the 3rd DVS point when using low nitrogen fertilizer
SLW_{max}	550	$kg\,ha^{-1}$	maximum specific leaf weight
SLW_{min}	250	$kg\,ha^{-1}$	minimum specific leaf weight
SLW_x	2.5	-	exponential slope of specific leaf weight to the developmental stage
$V_{ctopmax}$	103×10^{-6}	$mol(CO_2)m^{-2}s^{-1}$	maximum Rubisco capacity at the canopy top in Eq. (7)

Deleted: 600

Deleted: m^{-2}

Deleted: m^{-2}

Deleted: $\mu mol(CO2)m^{-2}$

3 Model Calibration

The model parameters were tuned to represent the observed phenology and seasonality of biomass development. Once calibration is complete, the model continues to simulate crop growth, which encompasses phenological development, carbon assimilation, assimilate partitioning, and crop yield. We conducted calibrations from various environmental conditions and soybean varieties documented in previous experimental studies as detailed in 3.1 and Table 3. The model calibration included parameterizing the dynamic biomass growth partitioning ratio for each organ, leaf senescence, and specific leaf weight denoted as P_{organ} during the developmental stage denoted as DVS. Other calibrations using the experimental dataset included the phenological stage, and the seed-pod ratio (SR). The crucial phenological stage (e.g. flowering and seed filling) was calculated as the average value of the reported values in the experimental dataset. MATCRO applies this crop growth module following the method by the school of de Wit, compares biomass growth with the observed values during developmental stages. Shifts in partitioning and growth patterns were identified and used as reference points in the parameterization.

Deleted: 3.1

Deleted: **Table 3**.

Deleted: Table 3

Formatted: Font: 10 pt, Not Bold

Deleted: proposed

Deleted: used in MACROS (Vries et al., 1989), and

3.1 Description of the site data for calibration

The calibration process used experimental datasets from previous studies collected from field experiments across six different sites in four countries: Frederico Westphalen and Piracicaba (Brazil), Ya'an (China), Champaign (United States of America, US), Morioka and Tsukubamirai (Japan), as seen in Table 3. The soybean cultivars grown at these experimental sites represented different maturity groups. A variety of management practices related to water management and nutrients were utilized in the experiments. Nitrogen fertilizers were applied in most experiments, but soil mineral nitrogen at the Brazil and the US have provided adequate supply to support crop growth. Furthermore, there are different farming practices based on the across countries. Soybeans are planted between May and June in the United States, China, and Japan, while planting starts in October or November in Brazil. The experimental data also shown broad planting density in China and Japan, while soybeans are typically grown at higher planting densities in the United States and Brazil.

Deleted: Table 3.

Deleted: Table 3

Formatted: Font: Not Bold

Deleted: mineral nitrogen was used in the soil at sites in Brazil and the US

Weather data were derived from the records at the meteorological station nearest to the experimental site. The climates at the respective sites were as follows. The ranges of daily mean air temperatures during the growing season was 18-30ºC in Frederico Westphalen (Brazil), 19-31 ºC in Piracicaba (Brazil), 17-27 ºC in Tsukubamirai (Japan), 14-25 ºC in Morioka (Japan), 18-26 ºC in Ya'an (China), and 15-28 ºC in Champaign (US). The seasonal precipitation (mm) for the sites were 1669 mm in Frederico Westphalen (Brazil), 679 mm in Piracicaba (Brazil), 453 mm in Morioka (Japan), 865 mm in Tsukubamirai (Japan), 1012 mm in Ya'an (China), and 787 mm in Champaign (US). The amount of solar radiation also differed among the experimental sites where China received lowest solar radiation and Brazil received highest solar radiation during the experimental period (Supplementary file Figure S2). These data represent diverse climatic conditions in soybean-producing countries. The field data used for calibration were collected across multiple crop seasons, specifically from 2002, 2003 to 2007 and from 2013 to 2016. These time periods were expected to capture the current climatic and environmental variability.

Table 3. Information on field-experimental data of location, crop season, variety, maturity group, water management, and nitrogen fertilizer, as well as the number of experiments for calibrating glucose partitioning ratio and evaluating the soybean yield simulations.

Location	Crop season	Variety (RMG*)	Date of planting	Water management, Nitrogen fertilizer (g N m^{-2}), Plant Density (plant m^{-2})	Experiments (n)	Reference
Brazil (Frederico Westphalen)	2013	BRS284 (6)	Oct 1, 18; Nov 8, 25, Dec 12	Rainfed, 0, 26-28	5	(Battisti et al., 2017)
Brazil (Piracicaba)	2013-2014	BRS284 (6)	Oct 18, Nov 14 (2013); Jan 8 (2014)	Irrigated and Rainfed, 0, 16-37	6	(Battisti et al., 2017)
China (Ya'an)	2014	11 cultivars (5-8)	June 11	Irrigated, NA, 10	15	(Wu et al., 2019)
	2014-2016	Texuan13 (7), Jiuyuehang (5), Nandou12 (6)	June 15 (2014); June 18 (2015); June 18 (2016)		9	
United States (Champaign)	2002, 2004-2007	Pioneer93B15 (3)	June 1 (2002); May 28 (2004); May 25 (2005, 2006); May 22 (2007)	Rainfed, 0, 25 - 53	8	(Morgan et al., 2005; Ainsworth et al., 2007)
Japan (Tsukubamirai)	2013-2015	Enrei (2), Fukuyutaka (4), Ryuhou (2)	June 12, July 31 (2013); June 17, July 17 (2014); June 4, 30 (2015)	Rainfed, 25-27, 9.5	16	(Nakano et al., 2021)
Japan (Morioka)	2013-2016	Ryuhou (2)	May 13, 28 (2013); May 16, 30 (2014); May 5, 14, 25, 29 (2015); May 30, June 6, 27 (2016)	Rainfed, 25-30, 9.5	10	(Kumagai, 2018; Kumagai, 2021)

*Relative maturity group

3.2 Biomass partitioning and specific leaf weight

This model represents carbon assimilation by incorporating the carbon fraction in dry matter and glucose allocation to various plant organs. The glucose ratio for each organ is parameterized based on measurements of leaf weight, leaf senescence, stem weight, pod weight, and specific leaf weight across different developmental stages. To simulate glucose partitioning, we used Eq. (17)–(24) to fit the segmented linear models to the experimental dataset (Figure 2 and Figure 4) and used the parameter values as shown in Table 2, as this value is used to obtain the average value of soybean partitioning behaviour. The segmented linear models for glucose partitioning were manually determined by visual inspections of the plot. This approach was chosen due to the challenges of applying nonlinear optimization under multiple constraints. Breakpoints in the developmental stage were determined based on assumed growth characteristics, such as leaf development declines after the seed-filling stage, while pod formation starts after flowering. We assumed increasing trend of glucose allocation to leaf and shoot development during the early stage when data were unavailable, with subsequent segments aligned with observed data trends. The calibrated glucose partitioning ratio varied across the varieties and environmental conditions and was derived by converting biomass growth into glucose allocation as outlined in Eqs. (11)–(16).

The parameterization reflected the observation data, as well as the linear growth of leaves and pods during the developmental stages. It was utilized for seed-pod ratio and phenology parameterization. The dashed lines in Figure 2 and 3

Deleted: S1
Formatted ... [1]
Formatted ... [2]
Formatted ... [3]
Deleted: Section Break (Continuous)
Formatted ... [4]
Formatted ... [6]
Inserted Cells ... [5]
Deleted: -2014
Formatted ... [7]
Formatted ... [8]
Formatted ... [9]
Formatted ... [10]
Formatted ... [11]
Deleted: 2017a)
Formatted ... [12]
Formatted ... [13]
Formatted ... [14]
Formatted ... [15]
Deleted: 2017a)
Deleted: 15
Formatted ... [16]
Formatted ... [17]
Formatted ... [18]
Formatted ... [19]
Split Cells ... [22]
Formatted ... [20]
Formatted ... [21]
Formatted ... [23]
Inserted Cells ... [27]
Formatted ... [24]
Formatted ... [25]
Formatted ... [26]
Formatted ... [28]
Formatted ... [29]
Formatted ... [30]
Formatted ... [31]
Formatted ... [32]
Deleted: 12
Formatted ... [33]
Formatted ... [34]
Formatted ... [35]
Formatted ... [36]
Formatted ... [37]
Deleted: ¶
Deleted: … ... [38]
Deleted: Figure 2
Deleted: Figure 4…) and used the parameter values as sho ... [41]
Formatted ... [39]
Formatted ... [40]
Deleted: Table 2…, as this value is used to obtain the aver ... [43]
Formatted ... [42]
Formatted ... [44]
Deleted: The dataset from Morioka (Japan) was not includ ... [45]

indicate the estimated flowering and seed filling stages by calculating the average phenology time in all experimental datasets. The independent dataset was used for evaluating the calibrated model at the point-scale level. After removing the calibration data, the simulated yield at the site scale showed a correlation coefficient of 0.68 and significancy (p value < 0.001) with observed data (Supplementary file Figure S3). This agreement is also applied for the aboveground biomass weight, pod weight, and leaf area index with correlation coefficient of 0.60−0.90.

Carbon assimilation primarily occurs with subsequent allocation to other parts of the plant. Compared with varieties from other sites, the soybean varieties observed in the experimental dataset from Tsukubamirai (Japan) tended to have lower partitioning to the stem during the vegetative stage. The ratio of glucose to leaves in Sichuan (China) was unexpectedly high near maturity in 2016, resulting in partitioning to pods at a low level due to low temperature and drought conditions. The storage organ biomass increases in the reproductive stage to produce pods and seeds, whereas the shoot will senesce at the end of the maturity period. Hence, yield is estimated using seed weight (as determined by the storage organ weight) and the parameterized seed-pod ratio. Pod partitioning in Champaign (US) tended to occur early in pod initiation in early maturation varieties, and the dry weight of pods before the seed filling stage is relatively high (Kawasaki et al., 2018). Early pod initiation has also been observed in the Ryuhou variety in Tsukubamirai in 2013 (Nakano et al., 2021).

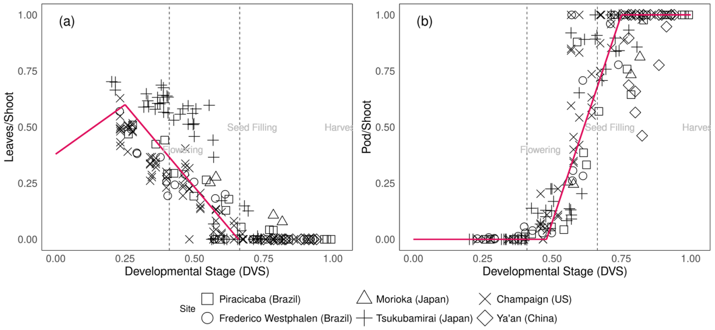
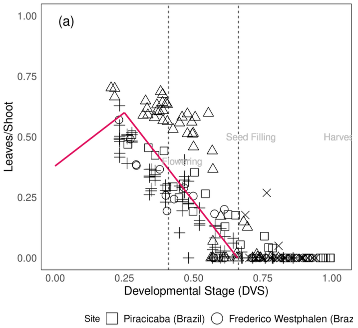


Figure 2. Glucose partitioning ratio to leaves (a) and pod (b) compared with the shoot during the developmental stage (DVS = 0 - 1) in the experimental sites shown by shaped points (square: Piracicaba, circle: Frederico Westphalen, triangle: Morioka, plus: Tsukubamirai, cross: Champaign, diamond: Ya'an. The red lines are the segmented lines used for glucose partitioning in MATCRO-Soy. The dashed line marks the averaged flowering, seed filling, and harvest time from the experimental datasets.

Deleted: in all experimental datasets



Deleted:

Deleted: plus

Deleted: cross

Deleted: fitted model of

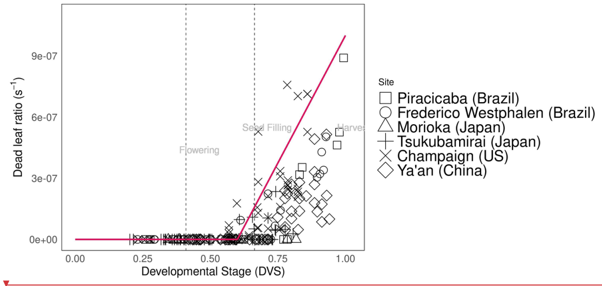


Figure 3. Dead leaf ratio (s-1) during the developmental stage (DVS = 0 - 1). Similar with Figure 2.

The dead leaf ratio parameter in Figure 3 shows the degree of leaf senescence after the seed filling stage due to the leaf process. The dead leaf ratio is calculated from the amount of leaf loss observed during the growing season. The specific leaf weight (SLW) is a significant parameter in crop growth parameterization and has been calibrated to follow the observation data pattern shown in Figure 4. We used the measured leaf weight and leaf area index data from the experimental datasets described in 2.4 and Eq. (23) to calculate the ratio of leaf weight to leaf area (SLW) during different phenological stages. These ratios change over time with distinct values as they vary across different growing seasons and cultivars (Thompson et al., 1996; Slattery et al., 2017). In the figure, SLW from Champaign (US) was excluded because of discrepancies in the timing of the measurements in leaf area and leaf weight biomass. While the specific leaf weight varied among the sites, we fit the model of SLW assuming a saturating exponential function of the developmental stage (red line in Figure 4). This pattern aligns well with the biological process as SLW initially increases due to rapid biomass accumulation but saturates as leaves mature.

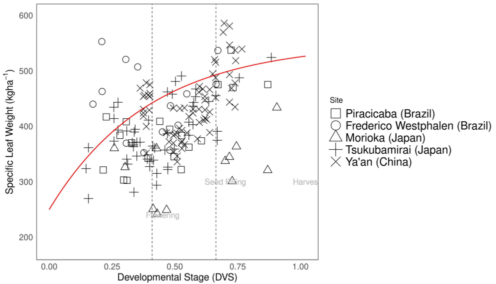
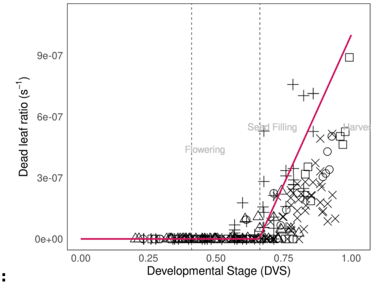


Figure 4. Specific leaf weight (kg ha^{-1}) during the developmental stage (DVS = 0 - 1). Similar with Figure 2.

4 Model Evaluation Setup

MATCRO was developed in FORTRAN and coupled with the global climate models output, simulated at a spatial resolution of 0.5° × 0.5° and hourly-daily temporal resolution. The output of the model is gridded crop yield (kg ha^{-1}) as stored in netCDF file format in a global map with one harvest simulated per year. We perform the model evaluation for global,



Deleted:

Formatted: Font color: Auto, English (UK)

Deleted: Figure 3

Formatted: Font: Not Bold

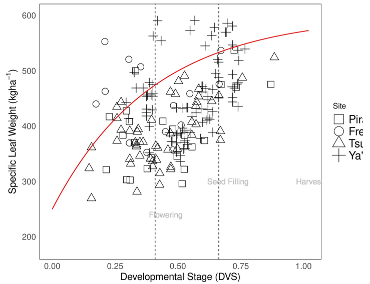
Deleted: Figure 4

Formatted: Font: Not Bold

Deleted: Figure 4

Formatted: Font: Not Bold

Formatted: Font color: Auto



Deleted:

Formatted: Font color: Auto

Deleted: kgha

Formatted: Superscript

Deleted: -time

Deleted: in one

country, and grid cell levels for 34 years (1981–2014) at 0.5° spatial resolution and yearly harvested yield output. The accuracy of the simulated yield was assessed using reference global and country-level data from the Food and Agriculture Organization (FAOSTAT, 2023), while the grid cell level yield was compared with the Global Dataset of Historical Yield (GDHY) data which is derived from statistical records, FAO data, and remote sensing data (Iizumi, 2019).

Deleted:

Deleted:

4.1 Simulation settings and data inputs

The parameters were set as shown in Table 4, covering the period of the sowing year from 1980 to 2014, with a various planting time across different regions. This model incorporated global daily climate data (86400 s) as input data. While the simulation framework was inherited from the established MATCRO-Rice v.1 (Masutomi et al. 2016b), several modifications were made to enhance its applicability at a global scale. Notably, the temporal resolution was adjusted from half-hourly (1800 s) to hourly (3600 s), allowing the model to maintain consistency in capturing critical processes such as diurnal variations in photosynthesis and transpiration, while optimizing computational efficiency. These adjustments ensured that the model remained suitable for large-scale simulations while preserving essential physiological processes.

Deleted: Table 4

Formatted: Font: Not Bold

The model simulates soybean yield using input data as described in Table 5. It uses global input data as follows: crop calendar from the Global Gridded Crop Model Intercomparison (GGCMI), which separates the rainfed and irrigated systems, atmospheric CO_2 and climate data from the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) that provides bias-adjusted climate input data for historical data (GSWP3-W5E5 v2.0), soil classification from the Harmonized World Soil Database (HWSD v1.2), and nitrogen fertilization for C3 fixing crops of the ISIMIP, which is derived from the land use dataset (Hurtt et al., 2020). We use ISIMIP bias-adjusted data to maintain uniformity in the climate impact data across sectors and scales in their framework. This dataset, which is provided by ISIMIP, has a spatial resolution of 0.5 °. To determine the growing degree days for maturity, we considered the phenological maturity time from the GGCMI crop calendar for harvest time and global ISIMIP climate data over 10 years (2000-2010) to capture the variability shifts in the current evaluation years.

Deleted: Table 5

Formatted: Font: Not Bold

Table 4. Parameter settings for simulation

Deleted: ¶

Variable	Value	Unit	Description
$Year_{sow}$	varied	Year	year of sowing day
DOY_{sow}	varied	Day of Year (DOY)	day of year of sowing day
RES_{time}	3600	s	time resolution for simulation
$RES_{climate}$	86400	s	time resolution for climate forcing data
$RES_{we/ns}$	0.5	degree	spatial resolution north to south or west to east
Soil layer	5.0	-	number of simulated soil layer to calculate soil water content
WSL	1.0	-	soil water content at emergence
W_{leaf0}	1.0	kg ha^{-1}	dry weight of leaf at emergence
W_{stem0}	1.0	kg ha^{-1}	dry weight of stem at emergence
W_{root0}	1.0	kg ha^{-1}	dry weight of root at emergence
W_{glu0}	0.5	kg ha^{-1}	dry weight of glucose reserve at emergence
Z_a	3.0	m	reference height at which wind speed is observed
Z_{max}	4.0	m	depth of soil layer
Zt	0.05	m	depth of topsoil layer
Z_b	2.0	m	depth from the soil surface to the upper bound of the most bottom layer of soil

Table 5. Data input for MATCRO simulation.

Variable	Unit	Data source	Spatial Resolution
	Daily time-step		
Precipitation	kg m^{-2}s^{-1}		0.5° × 0.5°

Deleted: kgm

Near-surface specific humidity	kg kg^{-1}	GSWP3–W5E5 (Kim, 2017; Cuchi et al., 2020; Lange, 2019; Lange et al., 2021)	0.5° × 0.5°
Maximum, minimum, and mean temperature	Kelvin		0.5° × 0.5°
Surface downwelling shortwave radiation	W m^{-2}		0.5° × 0.5°
Near-surface wind speed	m s^{-1}		0.5° × 0.5°
Surface air pressure	Pa		0.5° × 0.5°
	Yearly time-step		
Atmospheric CO2 concentration	ppm	ISIMIP (Büchner and Reyer, 2022)	-
Nitrogen fertilizer	kg ha^{-1}	ISIMIP (Volkholz and Ostberg, 2022)	0.5° × 0.5°
	Constants		
Latitude and longitude	°	-	-
Agricultural management	Irrigated or rainfed	MIRCA2000 (Portmann et al., 2010)	0.5° × 0.5°
Sowing time, Harvest time	DOY	GGCMI (Jägermeyr et al., 2021)	0.5° × 0.5°
Growing degree days for harvest time	°C days	Parameterized in this study	0.5° × 0.5°
Soil type	-	HWSD (Volkholz and Müller, 2020)	0.5° × 0.5°

4.2 Global yield evaluation methods

In this study, we assessed the statistical relationship between simulated yields and reference data using common metrics of Pearson correlation coefficient (corr) in Eq (28) with the significance levels (p-values), agreement between the simulation and observation using root mean square error (RMSE) in Eq. (29), and bias in Eq. (30) for the time-series yield data.

$$corr = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2(Y_i-\bar{Y})^2}} \qquad (28)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2} \qquad (29)$$

$$relative\ bias = \frac{1}{n}\sum_{i=1}^{n}|X_i - Y_i| \times \frac{1}{\bar{Y}} \qquad (30)$$

where X_i and Y_i indicated simulated and observed values in each measurement, while $\bar{X}$ and $\bar{Y}$ denotes the mean of simulated and observed values for the harvested year annually. The i and n shows the i-th data point and total number of data, respectively. We use n = 34 years for global-scale data, while output after calibration is evaluated in point-scale using n ranged from 14-122 of the available experimental datasets.

Detrended yield represents the time-series yield data for both simulated and observed values after removing the linear trend by subtracting the slope and intercept of the fitted linear regression (long-term yield trend). This approach enables the separation of short-term yield fluctuations from systemic long-term shifts. Yield fluctuations for the long-term and detrended data were evaluated separately using mean squared deviation (MSD) and its component to provide a clear interpretation of the model evaluation error (Gauch et al., 2003; Kobayashi and Salam, 2000) in Eq. (31)

$$MSD_y = SB_y + SDSD_y + LCS_y \qquad (31)$$

Mean squared deviation (MSD_y) is the square of RMSE for each long-term yield trend or detrended yield. Its components included mean squared bias (SB_y), difference in the magnitude of fluctuation namely squared difference between standard deviations ($SDSD_y$), and the lack of positive correlation weighted by the standard deviations (LCS_y) as proposed by Kobayashi and Salam (2000) calculated in Eq (32)–(37) below:

$$SB_y = (\bar{X} - \bar{Y})^2 \qquad (32)$$

$$SDSD_y = (SD_X - SD_Y)^2 \qquad (33)$$

Deleted:

Deleted:

Deleted: Julian day

Deleted: significancy

Deleted: Furthermore, we evaluated

Deleted: fluctuation on

Deleted: detrended

Deleted:

Deleted: standard

Deleted: more

Deleted: Detrended yield is calculated based on the value after yield is reduced by its long-term trend.

Deleted:

Deleted:

$$SD_X = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad (34)$$

$$SD_Y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (35)$$

$$LCS_y = SD_X SD_Y (1 - corr) \quad (36)$$

Higher SB_y, $SDSD_y$, and LCS_y indicate that model failed to simulate mean of the measurement, magnitude of fluctuation around the mean, and pattern of fluctuation across the n measurements, respectively, of the yield. SD_X and SD_Y denotes the standard deviation of simulated (X) and observed values (Y), while LCS_y depends on the correlation coefficient (corr).

5 Model Performance Evaluation

We calculated soybean yield in a global-scale map based on the gridded data of irrigated and rainfed area from MIRCA2000 dataset, which represents global agricultural land use around the year 2000 (Portmann et al., 2010), to get the actual yield value. We evaluated yield during the period of 1981-2014 as the MIRCA dataset was available within that period. The simulated yield at the country and global scales for regional comparison was determined by aggregating grid cell data to compute the mean soybean harvested area within each country grid as described below in Eq (37):

$$Yield_{region} = \frac{\sum_{i=1}^{n}[(Yield_{rf})_i(Area_{rf})_i + (Yield_{ir})_i(Area_{ir})_i]}{\sum_{i=1}^{n}[(Area_{rf})_i + (Area_{ir})_i]} \quad (37)$$

where $Yield_{region}$ is the aggregated yield at a given region (country or global-scale) in kgha^{-1} from the grid cell number (i) range from 1 to n (total number of grid cells in the region). The estimated yield under rainfed and irrigated conditions are denoted by $Yield_{rf}$ and $Yield_{ir}$, respectively. While the soybean rainfed and irrigated area (ha) used in the simulations are $Area_{rf}$ and $Area_{ir}$, respectively.

5.1 Model output yield as evaluated at the global and national scales

Figure 5a shows a time-series comparison from 1981 to 2014 between the global mean yields reported by FAOSTAT and those simulated by MATCRO-Soy. The results indicated that the model captures the upwards trend in yields over the period with smaller slope compared with the reported yield data. The correlation coefficient is 0.81, which is significant ($p < 0.01$). The errors were 298 kg ha^{-1} and 0.12 for the RMSE and relative bias, respectively. Notably, the simulated linear increase contributed to the higher correlation coefficient for the yield trends.

Figure 5b shows the comparison between the detrended global mean yield observed by FAOSTAT and the simulated value by MATCRO-Soy after removing the long-term linear trend across the study period. Detrended yield is the value after yield is reduced by its long-term trend from the original yield data. It isolates the variability primarily driven by climate fluctuations to evaluate interannual variability independent of long-term trends. However, it also removes longer-term signals (e.g. effect of technological improvements or increasing CO_2 concentrations). The correlation coefficient decreased to 0.446 ($p < 0.01$). The model reproduced the interannual variations well with an RMSE of 137 kg ha^{-1}. Specifically, according to observations, there were significant yield reductions in the years 1983, 1988, 2009, and 2012. Among these, the model successfully reproduced the yield reductions in three years (1983, 1988, and 2012), excluding 2009. These years are reported to have experienced severe droughts, and the model's ability to capture these events is noteworthy.

Deleted: 362

Deleted: 15

Deleted: correlation

Deleted: used in the

Deleted: was removed.

Deleted: 512, which was significant

Deleted: 124

Deleted: and a relative bias of 0.37

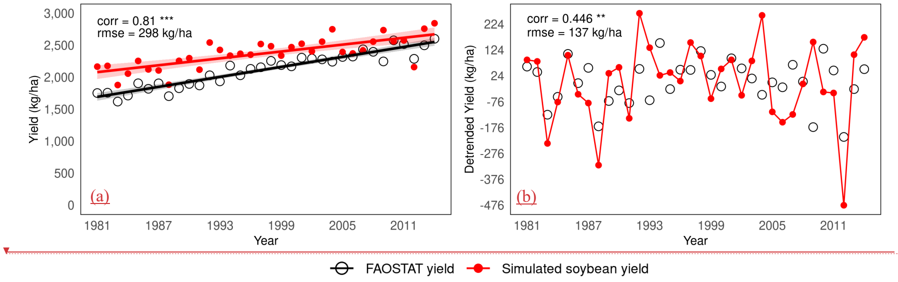
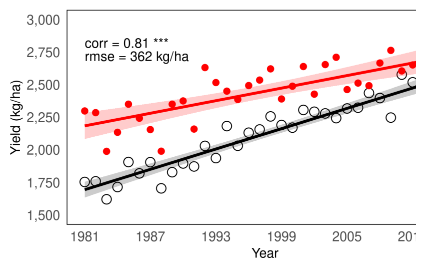


Figure 5. Time-series comparison between simulated yield by MATCRO-Soy and FAOSTAT reported yield data in global long-term trend (a), and detrended (b) yield during 1981-2014. The correlation for detrended yield is calculated after removing the linear trend. The symbols ***, **, and * denote p < 0.001, 0.01, and 0.05, respectively.

Deleted: *<object><object>*



We evaluated the model performance for 10 major soybean-producing countries, Argentina, Brazil, China, India, Paraguay, United States, Italy, Russia, Bolivia, and Canada, consisting of 96% of all global soybean production (based on total average production from 2012 to 2021 in FAOSTAT). Figure 6 compared the simulated country averaged yields and reported country averaged yields of FAOSTAT for 1981-2014 with the ellipsoid indicating the distribution of the simulated yield values within the 90% confidence range. The results indicate that the model reproduces the national average yield levels well in the top 10 producing countries, as indicated by a correlation coefficient of 0.519 (p < 0.001) and an RMSE of 1085 kg ha^{-1}. Significant correlation coefficients were observed for six countries (Argentina, Brazil, India, Italy, Paraguay, and the United States; see Supplementary file Figure S4 for further evaluation for these six countries). Focusing on the United States, Brazil, and Argentina, which account for 69% of global production, the model's accuracy showed a correlation coefficient of 0.645 (p < 0.001) and an RMSE of 916 kg ha^{-1}, where Brazil was underestimated. However, when all countries are considered, the correlation coefficient decreases to 0.291, although it remains statistically significant. These results demonstrate that the model performs reasonably well in capturing yield variations in major producing countries and achieves particularly lower bias in some countries (e.g. the United States, Italy, and Canada).

Deleted: Figure 6

Deleted: between

Formatted: Font: Not Bold

Deleted: 502

Deleted: 995

Deleted: S3

Deleted: top three producing countries (the

Deleted:),

Deleted: further improves, with

Deleted: 81

Deleted: 362

Deleted: .

Deleted: 263

Deleted: high accuracy in reproducing yields for

Deleted: with relatively high production levels.

Deleted: *<object>* ... [47]

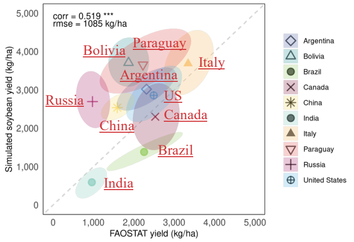


Figure 6. Comparison between simulated yield by MATCRO-Soy and FAOSTAT of the country mean yield during 1981-2014 in 10 major soybean producing countries. Ellipsoid shows 90% confidence range of annual yield.

A time series comparison of country averaged yields focusing on the major producing countries is shown in **Figure 7**. An evaluation of the long-term trend (Figure 7a) revealed that MATCRO-Soy effectively captured the increasing trend. Brazil demonstrated the strongest agreement, followed by Argentina at 0.62 and the United States at 0.64. For detrended yield (Figure

Deleted: **Error! Reference source not found.**

Deleted: Figure 7

Formatted: Font: Not Bold

Deleted: 73

Deleted: 62

Formatted: Font: Not Bold

7b), the interannual variability in Paraguay presented the highest correlation coefficient at 0.61, followed that in the United States at 0.57 and that in Brazil at 0.49. On the other hand, the lowest correlation was observed for China at 0.18 and Bolivia at -0.32. These findings suggest that the model tends to perform with greater accuracy for countries with higher production levels, even in time series comparisons at the national level.

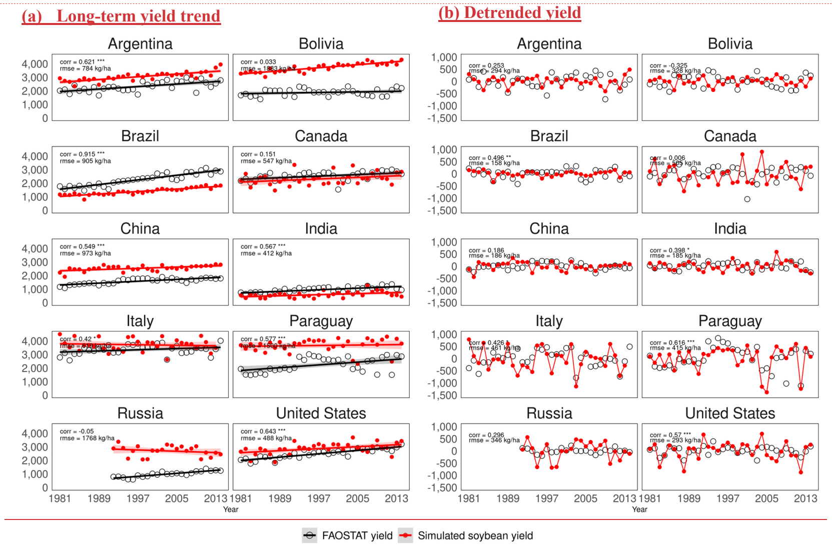


Figure 7. Time-series comparison between simulated yield by MATCRO-Soy (red circle) and FAOSTAT yield (open circle) in 10 top soybean producer countries during 1981-2014 for long-term yield trend shown by solid line (a) and detrended yield after removing the linear trend (b) in kg ha^{-1}. The correlation and RMSE based on yield (a) and detrended yield (b) data. The symbols ***, **, and * denote p < 0.001, 0.01, and 0.05, respectively. The shading near solid line is the standard error with confidence interval of 95%.

5.2 Temporal trends and variability

Model performance was further assessed with the mean squared deviation (MSD) components for the yield and separated by yield, long-term yield trend, and detrended yield for both the global (Supplementary file Table S1) and country scales (Supplementary files Table S2, S3, and S4). We separated the MSD into squared bias (SB), the sum of the difference in standard deviation (SDSD), and the lack of positive correlation (LCS), which reflect errors in mean yield, magnitude of yield variability, and pattern of year-to-year fluctuations, respectively. The greatest contributor to the error at the global scale was the difference of mean yield (SB) for about 71 and 77 % of total MSD for the yield and detrended yield, respectively (Supplementary file Table S1).

Figure 8 presents MSD components in the top six soybean-producing countries. SB was the primary source of error in most countries at the country-level. The highest MSD in Paraguay was largely driven by SB, with a notable contribution from LCS. It indicates that the model simulated the variability well but poorly captured the mean yield. The low MSD in the United States was also driven by SB, but LCS also contributed meaningfully to the year-to-year variability. Meanwhile, LCS was the

Deleted: Figure 7

Deleted: Brazil

Deleted: 79

Deleted: 6

Deleted: Paraguay

Deleted: 562

Deleted: 136

Deleted: 107

Deleted: *<object><object>*

Deleted: ¶ ... [48]

Deleted:).

Deleted: Figure 8 shows the source of error based on the MSD components of…

Deleted:) in the top 6 soybean-producing countries. SBs are the primary source of error in countries with high MSDs: Argent ... [49]

Deleted: Figure 8

Formatted: Font: Not Bold

Deleted: minor

Deleted: In contrast, the lowest MSD in Brazil was largely driven by SDSD and LCS. The SDSD is the primary contributor to ... [50]

Deleted: effectively

Deleted: mean yield

Deleted: trends.

greatest contributor of yield error in Canada and Italy (Supplementary file Table S2) due to a pronounced discrepancy in the simulated interannual variability. SDSD contributed only to Brazil, and the model underestimated the mean yield and the deviation in this country. These results highlighted that the mean yield bias dominates the source of error at global and country levels, while LCS and SDSD contributed notably in specific regions where the model failed to capture the variability or the temporal pattern.

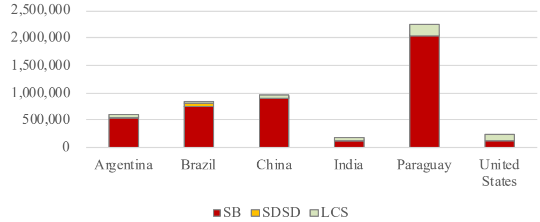


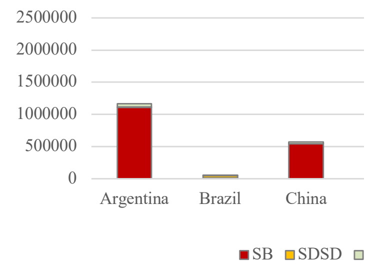
Figure 8. Mean squared deviation components of squared bias (SB), sum of difference in standard deviation (SDSD), lack of positive correlation (LCS) for yield error in top six soybean producing countries.

5.3 Model performance at the grid-cell level variation

We evaluated MATCRO-Soy at the grid-cell level, by comparing simulated yields with observed ones from Global Dataset of Historical Yield (GDHY) dataset by Iizumi (2019). Figure 9a and b show the simulated and observed yields averaged over 34 years, and Figure 9c shows relative bias between them. Figure 10 shows interannual correlation between simulated and observed yields for 34 years. The simulated yield was calculated for soybean-growing areas from the MIRCA2000 dataset, which offers broad spatial coverage where yield data for certain regions, including Canada, Russia, Australia, and many European and Asian countries, are missing in the GDHY dataset (Iizumi and Sakai, 2020). The density plot of the simulated yield showed more variability than did the GDHY data in Figure 9. However, both datasets exhibited a density peak of approximately 2,000 to 3,000 kg $ha^{-1}$and the simulated yield mostly overestimated the higher yield value. Figure 9 a, b, and c also show the distribution of simulated and observed yields.

The relative bias map (Figure 9c) highlights that overestimation was prominent in parts of South America (particularly Argentina), Russia, and China. In contrast, underestimation was observed in South Africa, India, and Brazil. Most of grid cells in Brazil show low yields, likely due to shorter growing periods in the input data compared to field experiment data. These results aligned with the trends observed at the national scale, which are influenced by the aggregation process. During aggregation, the national-scale results represented the average performance across all grid cells, weighted by the number of grids within each region. Most grids were within a relative bias of -0.2 to 0.2, accounting for 37 % of the total grid area. The grey area was found to be statistically insignificant. The density plot in simulated yield showed more variability compared to the GDHY data. However, both data exhibited the density peak around 3,000 kg ha^{-1} and simulated yield mostly overestimated the yield value. The correlation between the simulated yield and the GDHY dataset for interannual variation after removing the moving-average (Figure 10) reveals that 66 % of the grid cells are significantly correlated ($p < 0.05$).

Deleted: the model strength in simulating the mean yield in the top 2 major soybean producing countries (Brazil and the United States) with the largest soybean-growing areas.



Deleted:

Deleted: standard

Deleted: 6

Formatted: Font: Not Bold

Deleted: Figure 9

Formatted: Font: Not Bold

Deleted: Figure 9

Formatted: Font: Not Bold

Deleted: Figure 10

Deleted: Figure 9

Formatted: Font: Not Bold

Deleted: Figure 9

Formatted: Font: Not Bold

Deleted: Figure 9

Formatted: Font: Not Bold

Deleted:)

Deleted: , whereas

Deleted: India

Formatted: English (UK)

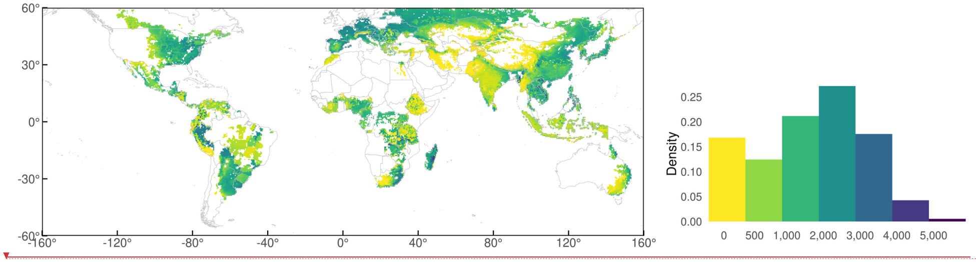
Deleted: 33

Deleted: Figure 10

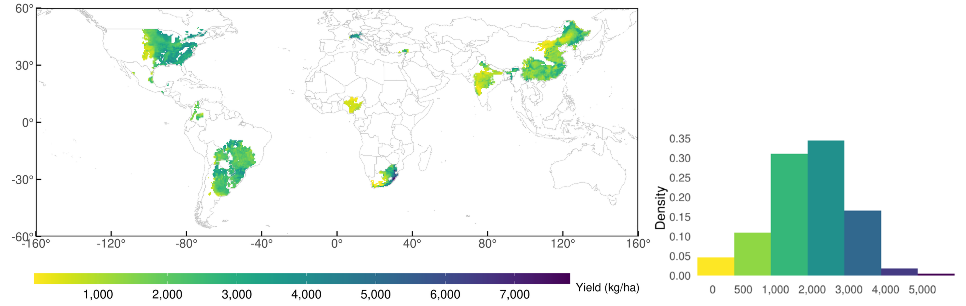
Deleted: 71

Formatted: Font: Not Bold

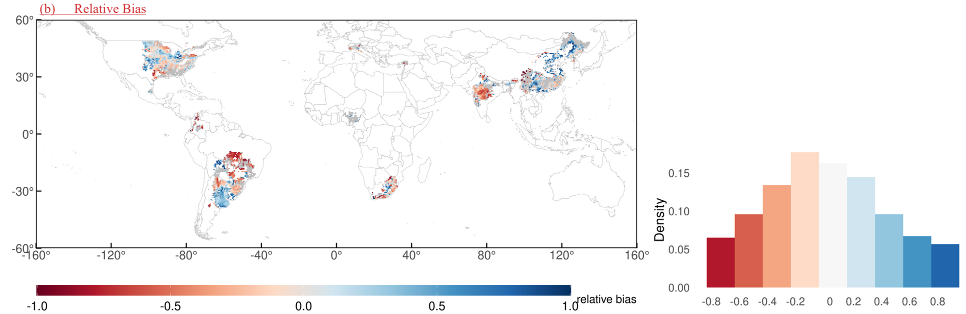
(a) MATCRO-Soy



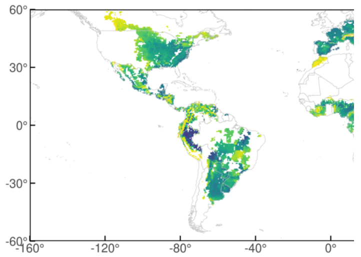
(c) GDHY



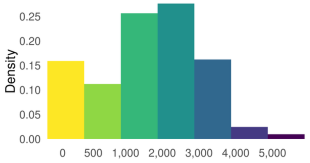
(b) Relative Bias



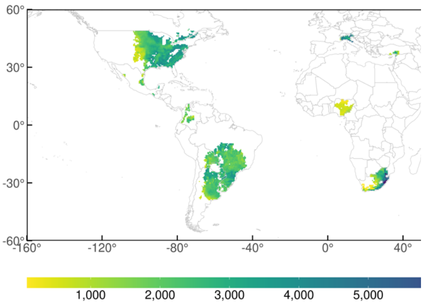
835 **Figure 9.** Global map of 34-year averaged (1981-2014) yield of GDHY dataset (a), simulated by MATCRO-Soy (b), and relative bias (c) with each density plot distribution. In figure c, grey colour depicts the correlation with no significance ($p > 0.05$) in the map.

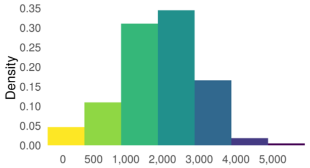


Deleted:



Deleted: *<object>*





... [51]

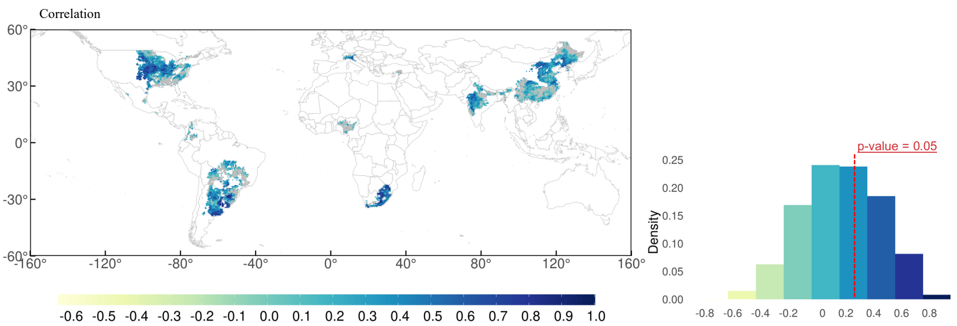


Figure 10. Time-series correlation between simulated and observed yield in 1981-2014 after removing trends from 5-year moving average (c). Grey colour depicts the correlation with no significance ($p > 0.05$) in the map while the red dashed line shows the border of $p = 0.05$ for the number of *n* year (34) in the density distribution plot.

5.4 Model performance at the leaf-level

We simulated the leaf-level variation in V_{cmax} for the United States (largest soybean producing country) at the site scale of the Champaign for the 2002 growing season using the global parameterization of MATCRO-Soy (Figure 11). These leaf-level simulated V_{cmax} values align closely with the observation data from Bernacchi et al. (2005) during the vegetative stage with some deviations during the flowering to seed-filling stages, as shown by the dashed line in the developmental stage of Figure 11. This alignment highlighted the ability of the model to represent essential photosynthetic processes influenced by leaf nitrogen content.

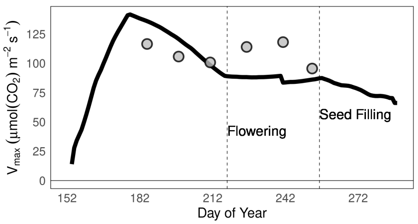


Figure 11. The maximum carboxylation capacity of Rubisco ($\mu mol\ (CO_2)\ m^{-2} s^{-1}$) during the growing period of simulation using MATCRO-Soy (black line) and observation data (grey dots) from Bernacchi et al. (2005) in Champaign (US) year 2002.

6 Discussions

6.1 Validation of MATCRO-Soy

In prior studies, soybean yield predictions often faced challenges in capturing crop responses to climatic variables. The MATCRO-Soy model effectively captures the linear trend in soybean yields, with higher accuracy for long-term trends (corr = 0.812) than for detrended yields (corr = 0.446), as shown in Figure 5. This result of the global detrended yield is improved

Deleted: ¶

Deleted: Figure 11

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Deleted: Figure 11

Deleted: $(CO2)m^{-2}$

Deleted: 81

Deleted: 512

Deleted: Figure 5

Formatted: Font: Not Bold

compared with that of benchmark studies conducted by Müller et al. (2017), indicating less variation among the process-based models based on its statistical correlation, where another crop model, PRYSBI2, reaches significant correlations of 0.57 ($p < 0.050$) if trends are not removed. However, the accuracy is enhanced when using site-specific parameters are used, as demonstrated in regional scale evaluations from previous studies, which were used for parameterization in this global simulation (Battisti et al., 2017; Kumagai, 2018, 2021; Morgan et al., 2005; Nakano et al., 2021; Wu et al., 2019). These studies have shown that integrating factors of cultivar differences, ensembles of multiple crop models, nitrogen content, and more accurate measurement method allows for a more reliable representation of local growing conditions and climate variability.

When examining the 10 largest soybean-producing countries, the model performance (Figure 6) has an RMSE of 1,085 kg ha^{-1} (average yield of 34 years), which is in reasonable agreement compared with the RMSE of another study using LPJ-GUESS coupled with biological nitrogen fixation (Ma et al., 2022) of approximately 800 kg ha^{-1} (average yield of 10 years). The grid-cell level evaluation simulated by MATCRO-Soy, as shown in Figure 9, revealed that 66% of the grid cells were significantly correlated ($p < 0.05$) with most grids falling within 0.2–0.6. These findings align with other studies that show that time-series correlations in GGCM simulated soybean yields range from 0.25 to 0.65 due to discrepancies in the benchmark studies (Müller et al., 2017). This correlation reflects the detrended values, which are useful for evaluating interannual variability and the model sensitivity to climate fluctuations. However, detrending removes important long-term signals related to genetic improvements, cultivar and management changes, or increased CO_2 effects.

The correlation values between yield and detrended yield in Figure 5 and Figure 6 indicate that the increased correlation in model performance was due to the long-term yield trend. MATCRO-Soy could capture the trend of increased atmospheric CO_2 and nitrogen fertilizer inputs, despite of the interannual variability in climate conditions. The MSD calculation revealed that the lack of positive correlation was the major contributor error in Canada and Italy within the 10 top soybean producing countries (Supplementary file Table S2). Both countries have small squared biases (SBs), suggesting that MATCRO-Soy accurately represents the average productivity despite of the inability to capture the variability or amplitude of the yield trend over time within the region. Factors such as changes in sowing date, land use, pest management, cultivar maturity group, and planting density may contribute to discrepancies in soybean yield under climate change (Battisti et al., 2018a; Marin et al., 2022). Hence, there is a need for improved parameterization to better represent the dynamics of yield variability in countries such as Canada and Italy.

The high yields in Argentina and Paraguay reflect the consistency of favourable growing conditions (Figure 7a), particularly the alignment of daily temperatures and seasonal precipitation with critical growth stages, suggesting that these regions are less susceptible to interannual variability along with the geographic locations to receive more radiation for photosynthesis sources. The comparison of simulations and observations at the grid-cell level (Figure 10) reveals weak correlations with no statistical significance in high-latitude countries with low number of grid cells (e.g., Canada and Russia). The models that lack sensitivity to daylength are observed to contribute to more uncertainty (Battisti et al., 2018b). Moreover, the low simulated yield in India, which has a hot climate characterized by high mean daily temperatures of 27–28 °C (Supplementary file Figure S5) and low soil moisture during the growing season, highlights the capacity of the model to capture regional climatic challenges that impact productivity. These climatic challenges likely exacerbate heat stress during critical phenological stages, such as flowering and pod development, leading to reduced yields (Sinclair, 1986; Egli and Bruening, 2004). The contrasting regions of high and low soybean yields underscore the ability of the model to capture the complex interplay between climate and crop yields across diverse agroecological zones.

6.2 Model strength and application

We developed MATCRO-Soy v.1, a process-based eco-physiological model that uses the Farquhar equation to simulate the leaf-level photosynthesis. The Farquhar equation is a widely recognized framework in plant physiology that simulates the biochemical mechanisms of photosynthesis by describing the relationships among light intensity, CO_2 assimilation, and

Deleted: 2017b

Deleted: Figure 6

Deleted: 0.9 ton

Formatted: Font: Not Bold

Deleted: comparable

Deleted: 0.8 ton

Deleted: Figure 9

Deleted: 71

Formatted: Font: Not Bold

Deleted: benchmark studies (Müller et al., 2017).

Deleted: Figure 5

Deleted: Figure 6

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Deleted: sum

Deleted: the differences in the standard deviation (SDSD)

Deleted: Brazil,

Deleted: in

Deleted: 2018;

Deleted: Brazil

Deleted: Figure 8

Deleted: Figure 7

Formatted: Font: Not Bold

Deleted: Figure 11

Deleted: Figure 10

Formatted: Font: Not Bold

Deleted: 2018)

Deleted: –

Deleted: S4

Deleted: model

Rubisco enzyme activity (Farquhar et al., 1980; Scafaro et al., 2023). Through the integration of this equation into a gridded global crop model, MATCRO-Soy enhances the simulation of soybean growth and productivity under environmental changes to atmospheric CO_2, temperature, and water scarcity. These factors are important for predicting and understanding the mechanism of the impact of climate change on productivity. The calibration of MATCRO-Soy successfully represented the response of soybean growth to a wide range of climatic conditions, resulting in reliable global yield simulations using a single parameterization. While simplification may introduce errors, global tuning effectively minimizes these discrepancies in specific regions as this similar result also shown by Smith et al. (2014).

Improving photosynthetic efficiency is one of the key improvements, particularly through enhancing stomatal conductance and modifying Rubisco, the enzyme responsible for carbon fixation (Xu et al., 2022). We used V_{cmax} as a photosynthetic parameter as it quantifies the Rubisco activity that is responsible for catalysing the conversion of carbon dioxide into organic compounds. The peak Rubisco activity observed during the reproductive stage corresponds with trends in specific leaf nitrogen and implicitly affected by the additional nitrogen fertilizer (La Menza et al., 2023). The consideration of nitrogen fixation is important as it is sensitive to adverse environmental conditions, flooding, water deficit, and inadequate temperatures, all of which reduce N_2 fixation (Santachiara et al., 2019).

The simulated yield, LAI, aboveground biomass, and pod biomass from MATCRO-Soy were further compared at the point-scale level with experimental datasets with distinct datasets used for each step of calibration and evaluation (**Table 3**) prior to global-scale evaluation (Supplementary file Figure S3). While point-scale simulations employed the unified global parameters, the results demonstrated reasonable agreement with a p value < 0.01 and a bias of 30–63 % for harvested yield, the seasonal leaf area index, aboveground biomass, and pod biomass. The highest bias was observed for the seasonal LAI, which aligns with the underestimation of V_{cmax} during critical growth stages. MATCRO-Soy can reproduce photosynthesis parameters comparable to those of the observation data in site-scale analysis with overestimation in the reproductive stage (Figure 11).

MATCRO-Soy effectively uses high-quality climate data, soil information, and nitrogen fertilizer data to capture biophysical processes involved in soybean growth and yield formation based on previous studies. Its flexibility in spatial resolution enables its application across various scales, from local studies to global assessments. Moreover, the structure of MATCRO is easily coupled with climate models and atmospheric CO_2 to increase the accuracy of yield predictions through high-quality data input. This adaptability also enables integration with other land models, making it a valuable tool in both ecological and agricultural research. MATCRO-Soy can be continuously refined with new data and plant physiological knowledge, ensuring that it remains robust and adaptable. This adaptability makes it a valuable for researchers and policy-makers working towards sustainable agriculture and global food security.

The strength of MATCRO-Soy lies in its ability to simulate key physiological processes of soybean growth (e.g. photosynthesis, phenology, and biomass partitioning), under varying climatic conditions. Its process-based structure allows for sensitivity analysis for further environmental impacts evaluation, such as effects of elevated CO_2 and temperature stress. The model has been shown to reasonably capture the temporal dynamics of yield formation. In addition to climatic factors, variations in yield may be attributed to technological advancements, shifts in agricultural practices, and changes in crop management strategies outside the scope of model can further improve the accuracy at the local scale. For example, including pest and crop interaction may enhance the model's capability to reflect local yield response to climate-driven pest dynamics (Chen and Mccarl, 2001). The integration of crop models with remote sensing data will enhance its capability for monitoring and predicting crop productivity at finer spatial scales (Basso et al., 2001). However, it is important to acknowledge the limitations of the model, particularly its ability to predict yield variations under extreme or rapidly changing climatic conditions. Continuous updates of the experimental dataset are necessary to maintain its relevance and accuracy in predicting future soybean yields.

Deleted: (

Formatted: Font: 10 pt

Formatted: Caption, Justified

Deleted: **Table 3**

Deleted: **Table 3**

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Deleted: S2

Formatted: Font: 10 pt

Deleted: 001

Deleted: 20–60

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt, Subscript

Formatted: Font: 10 pt

Deleted: Figure 11

Deleted: Furthermore, the integration of crop models with remote sensing data will open new possibilities for monitoring and predicting crop productivity at finer spatial scales (Basso et al., 2001). Climate change may shift favourable conditions for high yields in the United States or worsen the challenge of low yields in India for producing soybean yields, making it essential for the model to project these trends accurately for future agricultural planning.

Formatted: Font color: Auto

Deleted: that have not been considered in the model. The consideration of more detailed mechanisms of soybean growth can be considered for more accurate results as climate change affects the pest populations (Chen and Mccarl,

Formatted: Font color: Auto

6.3 Model challenges and future directions

In the evaluation process, it is important to recognize the interannual variability and spatial variability. There are many grid cells that have a low correlation (nonsignificant) of soybean yield between the simulated and observed values in Brazil when considered in each single cell (Figure 9), but the correlation at the national-scale level is high (Figure 7). This means that local climatic factors affect soybean yield in Brazil. However, MATCRO-Soy is able to recognize broader regional trends leading to its aim at representing yield behaviour. Figure 12 presents the relative RMSE (RMSE value compared with the observation value) between the simulation and GDHY datasets for the detrended yield at the grid-scale. High relative RMSE values are observed in some parts of Africa (particularly in Nigeria), the United States, India, and China. Lower relative RMSE values are evident in regions such as Brazil and Argentina. India and the United States show low RMSEs at the national-level, but some grid cells within both countries have higher relative RMSEs at the grid-cell level. Detailed information on the spatial variation in the error components contributors is provided in Supplementary File Figure S6 for the long-term yield trend and Supplementary File Figure S7 for the detrended yield. These findings highlighted that number of grid cells significantly influence model performance, with regions containing fewer grids being more sensitive to localized factors and spatial heterogeneity during aggregation. These emphasize the importance of considering spatial resolution and representation when evaluating model performance.

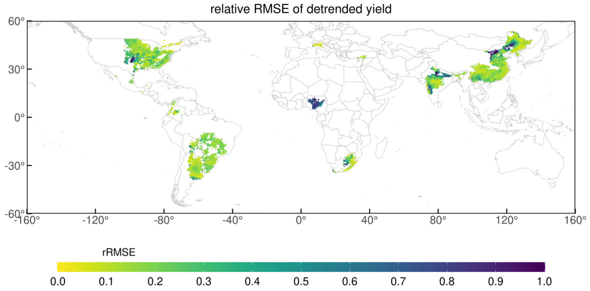


Figure 12. Relative RMSE calculation between simulated and observed yield for detrended yield in grid-cell level.

Uncertainty in MATCRO-Soy is reflected through the challenges in global-scale model evaluation related to the model assumptions of crop cultivars being homogenous globally and the upscaling parameters due to the lack of parameterization, making it is unrealistic to reproduce the variability at the regional-scale with very high accuracy (Müller et al., 2017; Zaehle and Friend, 2010). This uncertainty is notably pronounced in the global aggregation of yield simulations at the grid-cell scale. Global aggregation can escalate substantially for specific combinations of aggregation units, crop model limitations, and years (Porwollik et al., 2017). Future assessments of models and projections of crop yields will require careful consideration of the significant contrast between different aggregation approaches used for individual countries or regions. To address this, we used harmonized ISIMIP data to minimize methodological bias and emphasize the importance of flexible model development for reducing uncertainty (Yin, 2013).

We found a large underestimation in 2002, and overestimation in 2009 when comparing the soybean yield simulated using bias-corrected climate data was compared with FAO data (Figure 5). One possibility for these discrepancies in the interannual variability may be attributed to the influence of unaccounted extreme climatic events. Climatic events indicated by Oceanic Niño index, a three-month running mean of SST anomalies in the Niño 3.4 region, show that La Niña was present at the end of 2002 and that El Niño occurred at the end of 2009 (NOAA, 2024). Some regions within major soybean-producing countries

Deleted: Figure 7

Formatted: Font: Not Bold

Deleted: Figure 12

Formatted: Font: Not Bold

Deleted: South America

Deleted: Argentina and Paraguay

Deleted: . However, low RMSEs are found in Brazil

Deleted: although

Deleted: are found in some grid cells of

Deleted: United States

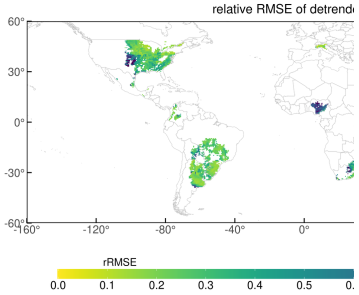
Deleted: S5

Deleted: S6

Deleted: the

Deleted: grids

Deleted:



Formatted: English (UK)

Deleted: Figure 5

Formatted: Font: Not Bold

are significantly affected by El Niño events, further influencing yield variability (Anderson et al., 2017; Iizumi et al., 2014). Another possibility for the interannual variation in MATCRO-Soy tends to overestimate the long-term yield trend because of the sensitive effect of the CO_2 concentration on the carbon assimilation module.

While nitrogen fixation and uptake are implicitly constrained by the SLN parameter, an approach of carbon costs economics explicitly represents the respiratory cost due to different nitrogen uptake pathways (Fisher et al., 2010). MATCRO-Soy simplified the nitrogen fixation mechanism, which may have contributed to yield overestimation in low nitrogen input countries (e.g. Bolivia and Russia). However, this model still presented relatively small bias in countries with high nitrogen fertilizer application (e.g. China), as well as in countries with low nitrogen fertilizer input (e.g. the United States). This highlights an opportunity for future model development to incorporate variable of respiratory costs in biological nitrogen fixation. While limited empirical data across cultivars, environments, and management systems poses a challenge at the global scale, addressing this would improve understanding of the physiological mechanisms under nitrogen-limited conditions.

The simulated yield increases throughout the year driven by the positive effects of increased atmospheric CO_2, a phenomenon known as the CO_2 fertilization effect, has been observed in studies by Long et al. (2005) and Sakurai et al. (2014). The CO_2 fertilization response may become a more prominent source of overestimation in future projections if the model overestimates the crop response to elevated CO_2. Compared with simulations using statistical radiation use efficiency (Ai and Hanasaki, 2023), process-based models have this tendency because of the greater effect of CO_2 on the photosynthesis process. Therefore, further investigation is needed into the CO_2 sensitivity of MATCRO-Soy and other process-based models, as the downregulation of photosynthesis under elevated CO_2 conditions has been observed in the measurements (Ainsworth et al., 2002; Zheng et al., 2019). This is especially important for adaptation studies, as reliable yield projections are critical for designing effective adaptation strategies under future climate scenarios.

Deleted: ,

Deleted: as

Deleted: (Ai and Hanasaki, 2023). This result is expected, as most of the simulated yield values were overestimated compared with the reference data, except for the yield in Canada, which was due to the low-temperature conditions.

MATCRO-Soy simulations showed that MSD component of SB was the dominant contributor in the global and country-level yield error. It indicates the bias was in the over or underestimation of average yield, rather than in variability of discrepancy in the year-to-year yield pattern (Figure 8). These results highlighted the model uncertainty in simulating mean yield for improvement in major soybean-producing countries with large cultivation areas. The model overestimated the long-term trend in some countries. Inaccurate representation of CO_2 fertilization effect may have contributed to the mean yield bias. Other possible contributing factors for the bias are the simplified assumption of no respiratory costs for symbiotic nitrogen fixation and insufficient representation of water stress responses. The accuracy of data input may partly reflect the inherent gap between field experiment data and national average yields, which are influenced by local farming practices. While these discrepancies between the country and global levels are insightful, it provides a valuable opportunity for model improvement.

Deleted: Figure 8

Formatted: Font: Not Bold

Comparative studies with other soybean models and refining the MATCRO-Soy on the basis of these findings will contribute to a more comprehensive understanding of its capabilities and limitations. Incorporating additional datasets will further enhance the model representation of real-world conditions. McCormick et al. (2021) suggested that integrating machine learning models could improve accuracy through the calibration process with numerous datasets. However, the use of mechanistic models embedded in MATCRO to simplify the process has proven valuable for understanding and predicting the impacts of environmental factors on agricultural systems. This model can be used to identify potential adaptation strategies, such as changes in planting dates or the development of new crop varieties, to mitigate the adverse effects of climate change on soybean production. However, the application of this model at the field-scale requires high-quality data input and local parameter data.

7 Conclusions

We utilized MATCRO which incorporates carbon assimilation modules based on the C3 photosynthesis of the Farquhar model, to simulate global soybean yield in terms of eco-physiological integrated gridded data inputs of climate, soil type, and

Deleted: utilize

nitrogen fertilizer. This study used experimental datasets and literature from previous studies to MATCRO-Soy to represent soybean growth under various environmental conditions. An evaluation of the global mean yield revealed a statistical correlation of 0.81 (p-value < 0.001) between the simulated and reported FAOSTAT data before the long-term yield trend was removed. The correlation value decreased after the long-term yield trend was removed, with a Pearson correlations of 0.45 ($p < 0.050$), 0.52 ($p < 0.001$), and 66 % grid cells statistically greater than the significant value ($p > 0.05$) over 34 years (1981-2014) for the global, top 10 countries, and grid cell levels, respectively. The model successfully captured long-term trends and interannual variability, demonstrating its capacity to reflect the impacts of climate factors. Moreover, MATCRO-Soy also modelled reasonable photosynthetic processes in site-scale study, which shows a strong ability to represent the temporal variation. This result highlights the model's reliability and adaptability as a tool for understanding soybean growth and yield dynamics.

Deleted: 51.2 %

Deleted: 50.2 %

Deleted: 71

While MATCRO-Soy presents a valuable framework for understanding the impacts of climate change on global soybean production, many localized factors that influence soybean yield due to the shifts in climate (e.g., pests and diseases) can lead to discrepancies in yield prediction. This highlights the need for high-quality data input. The integration of CO_2 dynamics in MATCRO enhances crop response modelling while providing the carbon fertilization effect in process-based models, warranting further investigation along with the effects of other greenhouse gases. The model may benefit for further refinement, particularly in its treatment of temperature extremes, transpirable soil water, and nitrogen uptake during the photosynthesis process. Integrating MATCRO with other environmental models would enhance its applicability in agricultural management, while emphasizing the necessity for field-scale calibration to improve the model's reliability. MATCRO-Soy provides an opportunity to estimate changes in global soybean production under future land-use or climate change scenarios to address the complexities of climate interactions with agricultural systems. Overall, the MATCRO-Soy has proven to be useful in understanding eco-physiological processes at both the global scale and the country and grid cell levels, providing valuable insights for agricultural management and climate change adaptation.

Code and data availability

This study used the model simulated by source code of MATCRO (Yusara et al, 2025) archived at https://doi.org/10.5281/zenodo.14881385.

Author Contributions

TK supervised this this study and acquired the funding. YM developed the model source code and supervised this study. AY adjusted the source code for the new parameterization and carried out the analysis and result's figures. RB, LA, YW, SN, and EK provided the data for model calibration. AY prepared the original draft with further editing by YM, TK, LA, YT, KK, RB, EK, SN, and YW. KK supervised the statistical analysis of MSD and provided result's figure. YM, TK, KK, YT, LA, RB, EK, SN, and YW reviewed it. Final manuscript was written by AY and approved by all authors.

Acknowledgements. This work was supported by JST SPRING, Grant Number JPMJSP2119, and was partially supported by a grant from The Tojuro Iijima Foundation, National Key Research and Development Program of China (2022YFD2300901), Environment Research and Technology Development Fund (JPMEERF21S12020 and JPMEERF20252002) of the Environmental Restoration and Conservation Agency (ERCA), and JSPS KAKENHI Grant Number JP23H00351. Author thanks all the contributors. Maki Nakaguki and Marin Nagata in Hokkaido University for their assistance in using the model. We acknowledge the data provided by Inter-Sectoral Impact Model Intercomparison Project.

Deleted: and

Deleted:).

Deleted: provided

Competing interests. There are no conflicts of interest to declare.

References

Adeboye, O. B., Schultz, B., Adeboye, A. P., Adekalu, K. O., and Osunbitan, J. A.: Application of the AquaCrop model in decision support for optimization of nitrogen fertilizer and water productivity of soybeans, Information Processing in Agriculture, 8, 419–436, https://doi.org/10.1016/j.inpa.2020.10.002, 2021.

Ai, Z. and Hanasaki, N.: Simulation of crop yield using the global hydrological model H08 (crp.v1), Geosci Model Dev, 16, 3275–3290, https://doi.org/10.5194/gmd-16-3275-2023, 2023.

Ainsworth, E. A., Davey, P. A., Bernacchi, C. J., Dermody, O. C., Heaton, E. A., Moore, D. J., Morgan, P. B., Naidu, S. L., Ra, H. S. Y., Zhu, X. G., Curtis, P. S., and Long, S. P.: A meta-analysis of elevated $[CO_2]$ effects on soybean (Glycine max) physiology, growth and yield, Glob Chang Biol, 8, 695–709, https://doi.org/10.1046/j.1365-2486.2002.00498.x, 2002.

Ainsworth, E. A., Rogers, A., Leakey, A. D. B., Heady, L. E., Gibon, Y., Stitt, M., and Schurr, U.: Does elevated atmospheric [CO2] alter diurnal C uptake and the balance of C and N metabolites in growing and fully expanded soybean leaves?, J Exp Bot, 58, 579–591, https://doi.org/10.1093/jxb/erl233, 2007.

Ainsworth, E. A., Serbin, S. P., Skoneczka, J. A., and Townsend, P. A.: Using leaf optical properties to detect ozone effects on foliar biochemistry, Photosynth Res, 119, 65–76, https://doi.org/10.1007/s11120-013-9837-y, 2014.

Anderson, W., Seager, R., Baethgen, W., and Cane, M.: Crop production variability in North and South America forced by life-cycles of the El Niño Southern Oscillation, Agric For Meteorol, 239, 151–165, https://doi.org/10.1016/j.agrformet.2017.03.008, 2017.

Ball, J. T.: An Analysis of Stomatal Conductance, Stanford University, CA, USA, 1988.

Basso, B., Ritchie, J. T., Pierce, F. J., Braga, R. P., and Jones, J. W.: Spatial validation of crop models for precision agriculture, Agric Syst, 68, 97–112, https://doi.org/10.1016/S0308-521X(00)00063-9, 2001.

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

Battisti, R., Sentelhas, P. C., Parker, P. S., Nendel, C., Câmara, G. M. D. S., Farias, J. R. B., and Basso, C. J.: Assessment of crop-management strategies to improve soybean resilience to climate change in Southern Brazil, Crop Pasture Sci, 69, 154–162, https://doi.org/10.1071/CP17293, 2018a.

Battisti, R., Sentelhas, P. C., and Boote, K. J.: Sensitivity and requirement of improvements of four soybean crop simulation models for climate change studies in Southern Brazil, Int J Biometeorol, 62, 823–832, https://doi.org/10.1007/s00484-017-1483-1, 2018b.

Moved (insertion) [1]

Moved (insertion) [2]

Moved (insertion) [3]

Moved (insertion) [4]

Moved (insertion) [5]

Formatted: Not Superscript/ Subscript

Deleted: 2017a

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

Bernacchi, C. J., Morgan, P. B., Ort, D. R., and Long, S. P.: The growth of soybean under free air [CO2] enrichment (FACE) stimulates photosynthesis while decreasing in vivo Rubisco capacity, Planta, 220, 434–446, https://doi.org/10.1007/s00425-004-1320-8, 2005.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M., and Swenson, S. C.: Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, J Geophys Res, 116, https://doi.org/10.1029/2010jg001593, 2011.

Boote, K. J., Jones, J. W., Hoogenboom, G., and Pickering, N. B.: The CROPGRO model for grain legumes, 1983, 99–128, https://doi.org/10.1007/978-94-017-3624-4_6, 1998.

Boote, K. J., Jones, J. W., White, J. W., Asseng, S., and Lizaso, J. I.: Putting mechanisms into crop production models, Plant Cell Environ, 36, 1658–1672, https://doi.org/10.1111/pce.12119, 2013.

Bouman, B. A. M., Kropff, M. J., Tuong, T. P., Wopereis, M. C. S., ten Berge, H. F. M., and van Laar, H. H.: ORYZA2000 : modeling lowland rice, IRRI, 235 pp., 2001.

Chen, C.-C. and Mccarl, B. A.: An investigation of the relationship between pesticide usage and climate change, ClimaticChange, 50, 475–487, 2001.

Collatz, G. J., Ball, J. T., Grivet, C., and Berry, J. A.: Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer, Agric For Meteorol, 54, 107–136, https://doi.org/10.1016/0168-1923(91)90002-8, 1991.

Cuddington, K., Fortin, M. J., Gerber, L. R., Hastings, A., Liebhold, A., O'connor, M., and Ray, C.: Process-based models are required to manage ecological systems in a changing world, Ecosphere, 4, https://doi.org/10.1890/ES12-00178.1, 2013.

Dai, Y., Dickinson, R. E., and Wang, Y.-P.: A Two-Big-Leaf Model for Canopy Temperature, Photosynthesis, and Stomatal Conductance, J Clim, 17, 2281–2299, 2004.

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K. J., Büchner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., Izaurralde, R. C., Mueller, N. D., Ray, D. K., Rosenzweig, C., Ruane, A. C., and Sheffield, J.: The Global Gridded Crop Model Intercomparison: Data and modeling protocols for Phase 1 (v1.0), Geosci Model Dev, 8, 261–277, https://doi.org/10.5194/gmd-8-261-2015, 2015.

Farquhar, G. D., Von Caemmerer, S., and Berry, J. A.: A Biochemical Model of Photosynthetic CO_2 Assimilation in Leaves of C 3 Species, Planta, 78–90 pp., 1980.

Fehr, W. R. and Caviness, C. E.: "Stages of soybean development" (1977). Special Report. 80., Iowa State University, 1–12, 1977.

Fisher, J. B., S. Sitch, Y. Malhi, R. A. Fisher, C. Huntingford, and S-Y. Tan.: Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation, Global Biogeochemical Cycles, 24, GB1014, 2010.

Formatted: Not Superscript/ Subscript

Moved down [6]: ., Arkebauer, T. J., Lindquist, J. L., Monzon,

Moved down [7]: J., Lindquist, J. L., Monzon, J. P., Knops, J. M. H., Graef, G., Scoby, D., Howard, R., Rees, J., Specht, J. E., and

Moved down [8]: E., and Grassini, P.: Decoupling between leaf nitrogen and radiation

Deleted: Büchner, M., Reyer, C. P. O.:: ISIMIP3b atmospheric composition input data (v1.1), ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.482153.1, 2022.¶
Cafaro La Menza, N

Deleted: use efficiency in vegetative and early reproductive stages in high-yielding soybean, J Exp Bot, 74, 352–363, https://doi.org/10.1093/jxb/erac408, 2023.¶

Moved up [1]: A.,

Deleted: Cucchi, M., Weedon, G. P., Amici,

Deleted: Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth Syst. Sci. Data, 12, 2097–2120, https://doi.org/10.5194/essd-12-2097-2020, 2020.¶

Deleted: CO 2

Fodor, N., Challinor, A., Droutsas, I., Ramirez-Villegas, J., Zabel, F., Koehler, A. K., and Foyer, C. H.: Integrating Plant Science and Crop Modeling: Assessment of the Impact of Climate Change on Soybean and Maize Production, https://doi.org/10.1093/pcp/pcx141, 1 November 2017.

García-Tejero, I. F., Durán-Zuazo, V. H., Muriel-Fernández, J. L., and Rodríguez-Pleguezuelo, C. R.: Water and Sustainable Agriculture, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-2091-6, 2011.

Gauch, H. G., Hwang, J. T. G., Fick, G. W., Gauch, H. G., and Fick, G. W.: Model Evaluation by Comparison of Model-Based Predictions and Measured Values, Agron J, 95, 1442–1446, 2003.

Hoogenboom, G.: Contribution of agrometeorology to the simulation of crop production and its applications, Agric For Meteorol, 103, 137–157, https://doi.org/10.1016/S0168-1923(00)00108-8, 2000.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850-2100 (LUH2) for CMIP6, Geosci Model Dev, 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Iizumi, T.: Global dataset of historical yields v1.2 and v1.3 aligned version, https://doi.org/10.1594/PANGAEA.909132, 2019.

Iizumi, T., Luo, J. J., Challinor, A. J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M. E., and Yamagata, T.: Impacts of El Niño Southern Oscillation on the global yields of major crops, Nat Commun, 5, https://doi.org/10.1038/ncomms4712, 2014.

Irmak, S., Odhiambo, L. O., Specht, J. E., and Djaman, K.: Hourly and daily single and basal evapotranspiration crop coefficients as a function of growing degree days, days after emergence, leaf area index, fractional green canopy cover, and plant phenology for soybean, Trans ASABE, 56, 1785–1803, 2013.

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J. A., Fuchs, K., Guarin, J. R., Heinke, J., Hoogenboom, G., Iizumi, T., Jain, A. K., Kelly, D., Khabarov, N., Lange, S., Lin, T. S., Liu, W., Mialyk, O., Minoli, S., Moyer, E. J., Okada, M., Phillips, M., Porter, C., Rabin, S. S., Scheer, C., Schneider, J. M., Schyns, J. F., Skalsky, R., Smerald, A., Stella, T., Stephens, H., Webber, H., Zabel, F., and Rosenzweig, C.: Climate impacts on global agriculture emerge earlier in new generation of climate and crop models, Nat Food, 2, 873–885, https://doi.org/10.1038/s43016-021-00400-y, 2021.

Jin, Z., Ainsworth, E. A., Leakey, A. D. B., and Lobell, D. B.: Increasing drought and diminishing benefits of elevated carbon dioxide for soybean yields across the US Midwest, Glob Chang Biol, 24, e522–e533, https://doi.org/10.1111/gcb.13946, 2018.

Kawasaki, Y., Katsura, K., and Shiraiwa, T.: Yield and dry matter dynamics of vegetative and reproductive organs in Japanese and US soybean cultivars, Plant Prod Sci, 21, 349–357, https://doi.org/10.1080/1343943X.2018.1512874, 2018.

Moved up [4]: L.,

Moved up [3]: B.,

Moved up [5]: S.,

Deleted: Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E., Dechow, R., Doro,

Deleted: Eckersten, H., Gaiser, T., Grosz,

Deleted: Heinlein, F., Kassie, B. T., Kersebaum, K. C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., Nendel, C., Priesack, E., Raynal, H., Roggero, P. P., Rötter, R. P., Siebert,

Deleted: Specka, X., Tao, F., Teixeira, E., Trombi, G., Wallach, D., Weihermüller, L., Yeluripati, J., and Ewert, F.: Impact of spatial soil and climate input data aggregation on regional Yield Simulations, PLoS One, 11, 1–23, https://doi.org/10.1371/journal.pone.0151782, 2016.¶

Deleted: Hoogenboom, G., White, J. W., and Messina, C. D.: From genome to crop: Integration through simulation modeling, in: Field Crops Research, 145–163, https://doi.org/10.1016/j.fcr.2004.07.014, 2004.¶
Hou, Z., Liu, B., and Kong, F.: Regulation of flowering and maturation in soybean, Adv Bot Res, 102, 43–75, https://doi.org/10.1016/BS.ABR.2022.02.007, 2022.¶

Deleted: J., Müller, C., Minoli, S., Ray, D., and Siebert, S., GGCMI Phase 3 crop calendar [data set], Zenodo, https://doi.org/10.5281/zenodo.5062513, 2021.¶
Jägermeyr

Kinose, Y., Masutomi, Y., Shiotsu, F., Hayashi, K., Ogawada, D., Gomez-Garcia, M., Matsumura, A., Takahashi, K., and Fukushi, K.: Impact assessment of climate change on the major rice cultivar ciherang in Indonesia, Journal of Agricultural Meteorology, 76, 19–28, https://doi.org/10.2480/agrmet.D-19-00045, 2020.

Kobayashi, K. and Salam, M. U.: Comparing simulated and measured values using mean squared deviation and its Components, Agron J, 92, 345–352, 2000.

Kothari, K., Battisti, R., Boote, K. J., Archontoulis, S. V., Confalone, A., Constantin, J., Cuadra, S. V., Debaeke, P., Faye, B., Grant, B., Hoogenboom, G., Jing, Q., van der Laan, M., Macena da Silva, F. A., Marin, F. R., Nehbandani, A., Nendel, C., Purcell, L. C., Qian, B., Ruane, A. C., Schoving, C., Silva, E. H. F. M., Smith, W., Soltani, A., Srivastava, A., Vieira, N. A., Slone, S., and Salmerón, M.: Are soybean models ready for climate change food impact assessments?, European Journal of Agronomy, 135, https://doi.org/10.1016/j.eja.2022.126482, 2022.

Kumagai, E.: Effect of early sowing on growth and yield of determinate and indeterminate soybean (Glycine max (L.) Merr) cultivars in a cool region of northern Japan, Journal of Agricultural Meteorology, 74, 18–28, https://doi.org/10.2480/agrmet.D-17-00009, 2018.

Kumagai, E.: Agronomic responses of soybean cultivars to narrow intra-row spacing in a cool region of northern Japan, Plant Prod Sci, 24, 29–40, https://doi.org/10.1080/1343943X.2020.1816137, 2021.

La Menza, N. C., Arkebauer, T. J., Lindquist, J. L., Monzon, J. P., Knops, J. M. H., Graef, G., Scoby, D., Howard, R., Rees, J., Specht, J. E., and Grassini, P.: Decoupling between leaf nitrogen and radiation-use efficiency in vegetative and early reproductive stages in high-yielding soybean, J Exp Bot, https://doi.org/10.1093/jxb/erac408, 2022.

Leung, F., Williams, K., Sitch, S., Tai, A. P. K., Wiltshire, A., Gornall, J., Ainsworth, E. A., Arkebauer, T., and Scoby, D.: Calibrating soybean parameters in JULES 5.0 from the US-Ne2/3 FLUXNET sites and the SoyFACE-O_3 experiment, Geosci Model Dev, 13, 6201–6213, https://doi.org/10.5194/gmd-13-6201-2020, 2020.

Long, S. P., Ainsworth, E. A., Leakey, A. D. B., and Morgan, P. B.: Global food insecurity. Treatment of major food crops with elevated carbon dioxide or ozone under large-scale fully open-air conditions suggests recent models may have overestimated future yields, in: Philosophical Transactions of the Royal Society B: Biological Sciences, 2011–2020, https://doi.org/10.1098/rstb.2005.1749, 2005.

Ma, J., Olin, S., Anthoni, P., Rabin, S. S., Bayer, A. D., Nyawira, S. S., and Arneth, A.: Modeling symbiotic biological nitrogen fixation in grain legumes globally with LPJ-GUESS, Geosci Model Dev, 15, 815–839, https://doi.org/10.5194/gmd-15-815-2022, 2022.

Marin, F. R., Ribeiro, R. V., and Marchiori, P. E. R.: How can crop modeling and plant physiology help to understand the plant responses to climate change? A case study with sugarcane, Theor Exp Plant Physiol, 26, 49–63, https://doi.org/10.1007/s40626-014-0006-2, 2014.

Marin, F. R., Zanon, A. J., Monzon, J. P., Andrade, J. F., Silva, E. H. F. M., Richter, G. L., Antolin, L. A. S., Ribeiro, B. S. M. R., Ribas, G. G., Battisti, R., Heinemann, A. B., and Grassini, P.: Protecting the Amazon forest and reducing global warming via agricultural intensification, Nat Sustain, 5, 1018–1026, https://doi.org/10.1038/s41893-022-00968-8, 2022.

Deleted: Kim, H: Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1), Data Integration and Analysis System (DIAS) [data set], https://doi.org/10.20783/DIAS.501, 2017.¶

Deleted: Simulated

Deleted: Measured Values Using Mean Squared Deviation

Moved up [2]: A.,

Moved (insertion) [6]

Moved (insertion) [7]

Deleted: Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.¶
Amici,

Deleted: Bellouin, N., Müller-Schmied, H., Hersbach, H., Buontempo, C., Cagnazzo, C: WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0), ISIMIP Repository [data set]. https://doi.org/10.48364/ISIMIP.342217, 2021.¶

Moved (insertion) [8]

Formatted: Subscript

Masutomi, Y., Ono, K., Mano, M., Maruyama, A., and Miyata, A.: A land surface model combined with a crop growth model for paddy rice (MATCRO-Rice v. 1) - Part 1: Model description, Geosci Model Dev, 9, 4133–4154, https://doi.org/10.5194/gmd-9-4133-2016, 2016a.

Masutomi, Y., Ono, K., Takimoto, T., Mano, M., Maruyama, A., and Miyata, A.: A land surface model combined with a crop growth model for paddy rice (MATCRO-Rice v. 1) - Part 2: Model validation, Geosci Model Dev, 9, 4155–4167, https://doi.org/10.5194/gmd-9-4155-2016, 2016b.

Masutomi, Y., Kinose, Y., Takimoto, T., Yonekura, T., Oue, H., and Kobayashi, K.: Ozone changes the linear relationship between photosynthesis and stomatal conductance and decreases water use efficiency in rice, Science of the Total Environment, 655, 1009–1016, https://doi.org/10.1016/j.scitotenv.2018.11.132, 2019.

McCormick, R. F., Truong, S. K., Rotundo, J., Gaspar, A. P., Kyle, D., Van Eeuwijk, F., and Messina, C. D.: Intercontinental prediction of soybean phenology via hybrid ensemble of knowledge-based and data-driven models, In Silico Plants, 3, https://doi.org/10.1093/insilicoplants/diab004, 2021.

Morgan, P. B., Bollero, G. A., Nelson, R. L., Dohleman, F. G., and Long, S. P.: Smaller than predicted increase in aboveground net primary production and yield of, Glob Chang Biol, 11, 1856–1865, https://doi.org/10.1111/j.1365-2486.2005.01017.x, 2005.

Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R. C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T. A. M., Ray, D. K., Reddy, A., Rosenzweig, C., Ruane, A. C., Sakurai, G., Schmid, E., Skalsky, R., Song, C. X., Wang, X., De Wit, A., and Yang, H.: Global gridded crop model evaluation: Benchmarking, skills, deficiencies and implications, Geosci Model Dev, 10, 1403–1422, https://doi.org/10.5194/gmd-10-1403-2017, 2017.

Nakano, S., Homma, K., and Shiraiwa, T.: Modeling biomass and yield production based on nitrogen accumulation in soybean grown in upland fields converted from paddy fields in Japan, Plant Prod Sci, 24, 440–453, https://doi.org/10.1080/1343943X.2021.1881409, 2021.

Portmann, F. T., Siebert, S., and Döll, P.: MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modelling, Global Biogeochemical Cycles, 24(1), http://doi.org/10.1029/2008GB003435, 2010.

Porwollik, V., Müller, C., Elliott, J., Chryssanthacopoulos, J., Iizumi, T., Ray, D. K., Ruane, A. C., Arneth, A., Balkovič, J., Ciais, P., Deryng, D., Folberth, C., Izaurralde, R. C., Jones, C. D., Khabarov, N., Lawrence, P. J., Liu, W., Pugh, T. A. M., Reddy, A., Sakurai, G., Schmid, E., Wang, X., de Wit, A., and Wu, X.: Spatial and temporal uncertainty of crop yield aggregations, European Journal of Agronomy, 88, 10–21, https://doi.org/10.1016/j.eja.2016.08.006, 2017.

De Pury, D. G. G. and Farquhar, G. D.: Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models, Plant Cell Environ, 20, 537–557, https://doi.org/10.1111/j.1365-3040.1997.00094.x, 1997.

Qiang, B., Zhou, W., Zhong, X., Fu, C., Cao, L., Zhang, Y., and Jin, X.: Effect of nitrogen application levels on the photosynthetic nitrogen distribution and use efficiency in leaves of soybean seedlings, https://doi.org/10.21203/rs.3.rs-2060928/v1, 2022.

Ray, D. K., Gerber, J. S., Macdonald, G. K., and West, P. C.: Climate variation explains a third of global crop yield variability, Nat Commun, 6, https://doi.org/10.1038/ncomms6989, 2015.

Ray, J. D. and Sinclair, T. R.: The effect of pot size on growth and transpiration of maize and soybean during water deficit stress, J Exp Bot, 49, 1381–1386, https://doi.org/10.1093/jxb/49.325.1381, 1998.

Ruane, A. C., Rosenzweig, C., Asseng, S., Boote, K. J., Elliott, J., Ewert, F., Jones, J. W., Martre, P., McDermid, S. P., Müller, C., Snyder, A., and Thorburn, P. J.: An AgMIP framework for improved agricultural representation in integrated assessment models, Environmental Research Letters, 12, https://doi.org/10.1088/1748-9326/aa8da6, 2017.

Sage, R. F.: Variation in the kcat of Rubisco in C3 and C4 plants and some implications for photosynthetic performance at high and low temperature, J Exp Bot, 53, 609–620, 2002.

Sakurai, G., Iizumi, T., Nishimori, M., and Yokozawa, M.: How much has the increase in atmospheric CO_2 directly affected past soybean production?, Sci Rep, 4, https://doi.org/10.1038/srep04978, 2014.

Santachiara, G., Salvagiotti, F., and Rotundo, J. L.: Nutritional and environmental effects on biological nitrogen fixation in soybean: A meta-analysis, Field Crops Res, 240, 106–115, https://doi.org/10.1016/j.fcr.2019.05.006, 2019.

Scafaro, A. P., Posch, B. C., Evans, J. R., Farquhar, G. D., and Atkin, O. K.: Rubisco deactivation and chloroplast electron transport rates co-limit photosynthesis above optimal leaf temperature in terrestrial plants, Nat Commun, 14, https://doi.org/10.1038/s41467-023-38496-4, 2023.

Sinclair, T. R.: Water and nitrogen limitations in soybean grain production I. Model development, Field Crops Res, 15, 125–141, https://doi.org/10.1016/0378-4290(86)90082-1, 1986.

Slattery, R. A., Vanloocke, A., Bernacchi, C. J., Zhu, X. G., and Ort, D. R.: Photosynthesis, light use efficiency, and yield of reduced-chlorophyll soybean mutants in field conditions, Front Plant Sci, 8, https://doi.org/10.3389/fpls.2017.00549, 2017.

Smith, B., Wärlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11, 2027–2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Stöckle, C. O. and Kemanian, A. R.: Can Crop Models Identify Critical Gaps in Genetics, Environment, and Management Interactions?, https://doi.org/10.3389/fpls.2020.00737, 12 June 2020.

Takata, K., Emori, S., and Watanabe, T.: Development of the minimal advanced treatments of surface interaction and runoff, Glob Planet Change, 38, 209–222, https://doi.org/10.1016/S0921-8181(03)00030-4, 2003.

Thies, J. E., Singleton, P. W., and Bohlool, B. B.: Phenology, growth, and yield of field-grown soybean and bush bean as a function of varying modes of N nutrition, Soil Biol Biochem, 27, 575–583, https://doi.org/10.1016/0038-0717(95)98634-Z, 1995.

Deleted: Rohatgi, A. WebPlotDigitizer (Version 5.2) [software], https://automeris.io, 2023.¶

Thompson, J. A., Schweitzer, E., and Nelson, R. L.: Association of specific leaf weight, an estimate of chlorophyll, and chlorophyll concentration with apparent photosynthesis in soybean, Phowsynthesis Research, Kluwer Academic Publishers, 1–10 pp., 1996.

Vries, F. W. T. P. de., Jansen, D. M., Berge, H. F. M. ten, and Bakema, A.: Simulation of ecophysiological processes of growth in several annual crops, Pudoc, 271 pp., 1989.

Wirth, S. B., Braun, J., Heinke, J., Ostberg, S., Rolinski, S., Schaphoff, S., Stenzel, F., Von Bloh, W., and Müller, C.: Biological nitrogen fixation of natural and agricultural vegetation simulated with LPJmL 5.7.9, https://doi.org/10.5194/egusphere-2023-2946, 2024.

Wu, X., Vuichard, N., Ciais, P., Viovy, N., De Noblet-Ducoudré, N., Wang, X., Magliulo, V., Wattenbach, M., Vitale, L., Di Tommasi, P., Moors, E. J., Jans, W., Elbers, J., Ceschia, E., Tallec, T., Bernhofer, C., Grünwald, T., Moureaux, C., Manise, T., Ligne, A., Cellier, P., Loubet, B., Larmanou, E., and Ripoche, D.: ORCHIDEE-CROP (v0), a new process-based agro-land surface model: Model description and evaluation over Europe, Geosci Model Dev, 9, 857–873, https://doi.org/10.5194/gmd-9-857-2016, 2016.

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

Xu, Q., Liu, X., Zhang, C., Du, W., Guan, Y., and Yang, W.: Insights into soybean with high photosynthetic efficiency, Adv Bot Res, 102, 121–151, https://doi.org/10.1016/BS.ABR.2022.02.019, 2022.

Yin, X.: Improving ecophysiological simulation models to predict the impact of elevated atmospheric CO2 concentration on crop productivity, https://doi.org/10.1093/aob/mct016, August 2013.

Yusara, A., Kato, T., Ainsworth, E., Rafael, B., Kumagai, E., Satoshi, N., Wu, Y., Tsusumi-Morita, Y., Kobayashi, K., Masutomi, Y.: An eco-physiological process-based model for soy yield (MATCRO-Soy v.1), Zenodo [code], https://doi.org/10.5281/zenodo.14881385, 2025.

Zaehle, S. and Friend, A. D.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates, Global Biogeochem Cycles, 24, https://doi.org/10.1029/2009GB003521, 2010.

Zheng, Y., Li, F., Hao, L., Yu, J., Guo, L., Zhou, H., Ma, C., Zhang, X., and Xu, M.: Elevated CO_2 concentration induces photosynthetic down-regulation with changes in leaf structure, non-structural carbohydrates and nitrogen content of soybean, BMC Plant Biol, 19, https://doi.org/10.1186/s12870-019-1788-9, 2019.

Deleted: Volkholz, J., Müller, C.: ISIMIP3 soil input data (v1.0), ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.942125, 2020.¶
Volkholz, J., Sebastian Ostberg, S.: ISIMIP3a N-fertilizer input data (v1.2), ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.311496.2, 2022.¶

Formatted: Not Superscript/ Subscript

Formatted: English (US)

Formatted: Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Formatted: Font color: Black

Deleted: ¶

Deleted: ¶

Page 11: [1] Formatted yusara.astrid 8/2/25 11:49:00 AM

Font: Bold

Page 11: [2] Formatted yusara.astrid 8/2/25 11:49:00 AM

Caption, Justified, Indent: Left: 0 cm, Hanging: 1 cm

Page 11: [3] Formatted yusara.astrid 8/2/25 11:49:00 AM

Font color: Auto

Page 11: [4] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [5] Inserted Cells yusara.astrid 8/2/25 11:49:00 AM

Inserted Cells

Page 11: [6] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [7] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [8] Formatted yusara.astrid 8/2/25 11:49:00 AM

Centered, Line spacing: single

Page 11: [9] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [10] Formatted yusara.astrid 8/2/25 11:49:00 AM

Right: 0.01 cm, Line spacing: single

Page 11: [11] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [12] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [13] Formatted yusara.astrid 8/2/25 11:49:00 AM

Centered, Line spacing: single

Page 11: [14] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [15] Formatted yusara.astrid 8/2/25 11:49:00 AM

Line spacing: single

Page 11: [16] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [17] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Centered, Line spacing: single

Page 11: [18] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [19] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [20] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Centered, Line spacing: single

Page 11: [21] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [22] Split Cells **yusara.astrid** **8/2/25 11:49:00 AM**

Split Cells

Page 11: [23] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Indent: Left: -0.24 cm, Line spacing: single

Page 11: [24] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [25] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Centered, Line spacing: single

Page 11: [26] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [27] Inserted Cells **yusara.astrid** **8/2/25 11:49:00 AM**

Inserted Cells

Page 11: [28] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [29] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Line spacing: single

Page 11: [30] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Centered, Line spacing: single

Page 11: [31] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Line spacing: single

Page 11: [32] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Line spacing: single

Page 11: [33] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Line spacing: single

Page 11: [34] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Centered, Line spacing: single

Page 11: [35] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Line spacing: single

Page 11: [36] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Line spacing: single

Page 11: [37] Formatted	yusara.astrid	8/2/25 11:49:00 AM

Line spacing: single

Page 11: [38] Deleted	yusara.astrid	8/3/25 1:40:00 AM

Page 11: [38] Deleted	yusara.astrid	8/3/25 1:40:00 AM

Page 11: [39] Formatted	yusara.astrid	8/3/25 2:48:00 AM

Font: Not Bold

Page 11: [40] Formatted	yusara.astrid	8/3/25 2:48:00 AM

Font: Not Bold, Do not check spelling or grammar

Page 11: [41] Deleted	yusara.astrid	8/3/25 1:57:00 AM

Page 11: [41] Deleted	yusara.astrid	8/3/25 1:57:00 AM

Page 11: [42] Formatted	yusara.astrid	8/3/25 2:48:00 AM

Font: Not Bold

Page 11: [43] Deleted **yusara.astrid** **8/3/25 1:57:00 AM**

Page 11: [43] Deleted **yusara.astrid** **8/3/25 1:57:00 AM**

Page 11: [44] Formatted **yusara.astrid** **8/2/25 11:49:00 AM**

Indent: First line: 0.5 cm

Page 11: [45] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**

Page 11: [45] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**

Page 11: [46] Formatted **yusara.astrid** **8/3/25 2:48:00 AM**

Font: Not Bold

Page 17: [47] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**

Page 18: [48] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**

Page 18: [49] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**

Page 18: [50] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**

Page 20: [51] Deleted **yusara.astrid** **8/2/25 11:49:00 AM**