Received: 27 August 2025 | Accepted: 9 January 2026

DOI: 10.1002/agj2.70309

Agronomy Journal

ORIGINAL ARTICLE

Biometry, Modeling, and Statistics

Calibration and evaluation of a carbon net primary productivity module based on the SIMPLE model

Michel Anderson Almeida Colmanetti[1] | **Henrique Oldoni**[1] | **Leandro Annibal**[1] | **Wagner Wolff**[2] | **Rodrigo Pereira Abou Rejaili**[2] | **Luis Gustavo Barioni**[1] | **Ivan Rodrigues de Almeida**[1] | **Robert P. Ewing**[2] | **Santiago Vianna Cuadra**[1]

[1]EMBRAPA Digital Agriculture, Campinas, São Paulo, Brazil

[2]Bayer Crop Science LLC, St. Louis, Missouri, USA

Correspondence
Michel Anderson Almeida Colmanetti, EMBRAPA Digital Agriculture, Campinas, SP 13083-886, Brazil.
Email: michelcolmanetti@gmail.com

Assigned to Associate Editor Yao Zhang.

Funding information
Carbon Project at Scale

Abstract

This study presents the carbon net primary productivity (CNPP) module, an implementation of the simple generic crop model (SIMPLE model) that we have extended to predict aboveground and belowground biomass production and carbon assimilation. The goal for this module is to predict biomass inputs at and below the soil surface, during and at the end of crop cycles, while integrated into a broader environmental carbon simulation framework, ProCarbon-Soil. CNPP was initially parameterized for soybean [*Glycine max* (L.) Merr.] and maize (*Zea mays* L.) using micrometeorological data, and for wheat (*Triticum aestivum* L.), bean (*Phaseolus vulgaris* L.), and perennial forage [*Urochloa* (syn. *Brachiaria*) *brizantha* (Hochst ex A. Rich.) Stapf cv. Marandu] using agrometeorological experimental data and/or data from the literature. Subsequently, calibrations for reference cultivars were performed, grouping cultivars by crop phenological characteristics and edaphoclimatic regions using farm-level data. CNPP accurately simulated leaf area index, evapotranspiration, and biomass dry matter production and allocation for soybean and maize when evaluated at the sites with micrometeorological data ($R^2 > 0.76$, Nash–Sutcliffe efficiency > 0.56, and relative root mean square error $< 38\%$ for all variables). Simulations for wheat, bean, and perennial forage exhibited lower performance owing to lower availability of yield data. Nonetheless, the resulting statistics support this module's efficacy in predicting crop productivity in major Brazilian agricultural areas. By employing a reduced and efficient parameter set, the CNPP module achieves enhanced performance and enables robust calibration across diverse crops, genotypes, and management schemes in multiple regions.

Abbreviations: AGB, aboveground biomass; AZD, active zone depth; BGB, belowground biomass; CNPP, carbon net primary productivity; MSD, maximum soil depth; DR, deep drainage; ET, evapotranspiration; HI, harvest index; LAI, leaf area index; MB, mean bias; PROCS, ProCarbon-Soil; RMSE, root mean square error; RRMSE, relative root mean square error; RUE, radiation use efficiency; RZD, root zone depth; RZDM, maximum root zone depth; SWAZ, soil water active zone; SWPZ, soil water passive zone; WPR, wheat producing region; WT, water transfer.

Plain Language Summary

The study presents the carbon net primary productivity (CNPP) module, a simplified, functional, and easily calibratable model for estimating crop biomass and carbon inputs to the soil. Based on the SIMPLE model, CNPP can be readily coupled with soil carbon models like ProCarbon-Soil, making it a valuable tool for estimating carbon sequestration in agricultural settings. The CNPP's design prioritizes a reduced parameter set to facilitate calibration across diverse crops and environments. The study aimed to describe the CNPP module, parameterize it for various crops, and evaluate its performance in Brazilian plantations. The study successfully demonstrated that the CNPP can provide reliable estimates of crop productivity in major Brazilian agricultural regions. While its performance is highly dependent on the quality of input data, the module's suitability for regional-scale analysis, even with limited on-farm data, the model has significant performance.

1 | INTRODUCTION

Increasing computational processing capacity and availability of experimental datasets worldwide (https://agmip.org/) have facilitated the development and use of increasingly complex models (e.g., DSSAT and APSIM). The simultaneous emergence of simpler, less glamorous models (e.g., SIMPLE and SSM-iCrop2) with reduced process and parameter sets is an important parallel development, as their easier parametrization and implementation suits them for cases where data are scarce or incomplete (Dadrasi et al., 2020; Soltani et al., 2020; Zhao et al., 2019).

Most biophysical crop models are designed to estimate biomass and yield, but even within that common mission their scope varies considerably. For example, DSSAT aimed to integrate knowledge about soil, climate, crops, and management to support better decisions about transferring production technology from one location to another (Jones et al., 2003). In contrast, APSIM was designed for accurate predictions of crop production in relation to climate, genotype, soil, and management factors, in order to optimize resources and address management issues in farming systems (Keating et al., 2003). Meanwhile, simpler models such as SIMPLE (Zhao et al., 2019) are more flexible in how they are implemented and calibrated, as they are used for less-studied, less commodified crops.

The CNPP module extends the foundational equations of the SIMPLE model (Zhao et al., 2019), building on its existing components to calculate carbon net primary productivity (CNPP). Its primary purpose is to serve as a driving component within the broader Conecta Pro Carbono simulation platform, where it is coupled to the ProCarbon-Soil (PROCS) model for simulating soil carbon stocks (Barioni et al., 2025). This framework was designed for commercial applications, specifically for estimating soil carbon stock within agricultural environments across various crop types, rotations, and management practices. This capability is particularly relevant given recent climate policy demands, which have heightened the necessity of using crop models to quantify carbon emissions and storage in managed croplands.

Similar to SIMPLE, the CNPP philosophy emphasizes a reduced set of equations and parameters, facilitating calibration across diverse crops, genotypes, environments, and management practices. However, integration within the PROCS framework necessitated some modifications. The most critical requirements were adjusting the allocation of carbon to belowground components, and implementing biomass deposition throughout the crop cycle as well as at harvest. Some modifications to the soil water module were also made.

The main objectives of this study are to present the CNPP module, designed specifically for integration into the PROCS environmental carbon simulation framework, and to calibrate and evaluate the CNPP module across Brazil's major agricultural regions for major grain crops. To achieve this, we address the following aims: (1) describe the CNPP module processes and parameters; (2) parameterize and evaluate the module for different crop types, including perennial forage; and (3) evaluate the module for simulating these crops in Brazilian farms.

2 | METHODS

2.1 | Module description

The CNPP is a crop module that simulates crop growth, biomass production, and harvested and residual biomass in a daily time step. It follows the SIMPLE model (Zhao et al., 2019) equations, particularly in its foundational equations for biomass production, and the influences of solar

radiation, temperature, and water on plant development. The CNPP's phenological framework, including harvest timing, aligns with SIMPLE by triggering events based on accumulated thermal time (TT reaching a crop-specific target). This approach was also employed by Crop Environmental Resources Synthesis (CERES)-wheat (Ritchie et al., 1985) for annual crops, but a different approach is needed for perennial crops. CNPP modifies SIMPLE's soil water balance calculation and water stress functions (Section 2.1.1) so it can consider the effect of root growth on plant water availability. Modifications were also made to biomass production and deposition of crop residues (Section 2.1.2). A schematic and representation of the CNPP, as well as this interaction with PROCS, is given in the Figure S1, as well as the parameters and input descriptions in Table S1. The subsequent sections will describe CNPP's major differences from SIMPLE.

2.1.1 | Soil water balance

To enhance hydrological realism, the CNPP module implements a two-compartment tipping bucket approach (Longo et al., 2021), replacing the simpler single-layer storage schema utilized by SIMPLE. The depth of the first layer is defined by the root zone depth (RZD) (Equations 1 and 2), and this layer holds the available water for plant uptake (near-surface soil water active zone [SWAZ]; Equation 4). Transfer between layers is given by Equation (3). The second layer (deeper soil water passive zone [SWPZ]; Equation 5) acts as a reservoir, allowing water to be held in a layer that can be progressively accessed by the plants as their roots grow. This is different from SIMPLE, which uses a fixed RZD as its maximum depth: in CNPP, RZD is simulated dynamically, reducing the potential available in the initial development of the plants and eventually reducing the RZD in conditions of high stress. This structural refinement permits the coupling of crop growth to root development and soil water balance, as considered in many crop growth models.

The RZD (mm day^{-1}) for a given date is calculated using Equation (1). The active zone depth (AZD; mm) is closely related to RZD, as described in Equation (2).

Core Ideas

- Carbon net primary productivity (CNPP) is a simple calibratable module for estimating crop growth and biomass.
- The module can be coupled with ProCarbon-Soil (PROCS) to estimate carbon sequestration.
- CNPP can be valuable for estimating crop growth in Brazilian agriculture.

$$RZD_t = \begin{cases} RZDGR, & t = 0 \\ min\left(\left(RZD_{t-1} + \min\left(\left(0.1 + f(Solar)_{t-1}\right), 1\right) \times RZDGR\right), RZDM\right), & t > 0 \end{cases} \quad (1)$$

$$AZD_t = \begin{cases} AZD_i, t = 1 \\ RZD_t, AZD_{t-1} \le RZD_t \\ AZD_{t-1}, AZD_{t-1} > RZD_t \end{cases} \quad (2)$$

where RZDGR is the root zone daily growth rate (mm), RZD_{t-1} is the root zone depth on day $t-1$, RZDM is the maximum RZD (mm), $f(Solar)_{t-1}$ is fraction of solar radiation intercepted on day $t-1$ (dimensionless), and AZD_t is the active zone depth (mm) on day t, with initial value AZD_i.

Soil water is computed daily for two different soil layers: (1) the near-surface SWAZ and (2) the deeper SWPZ. The upper layer "gains" water as the root depth increases; this water transfer (WT) is handled as follows:

$$WT_t = \frac{AZD_t - AZD_{t-1}}{MSD - AZD_{t-1}} \times SWPZ_{t-1} \quad (3)$$

$$SWAZ_t = \begin{cases} AZD_t \times AWC \times 0.7 + ERAIN - ET_t, t = 0 \\ ERAIN + SWAZ_{t-1} + WT_t - ET_t, t > 0 \end{cases} \quad (4)$$

$$SWPZ_t = \begin{cases} (MSD - AZD_t) \times AWC \times 0.7 - DR_t, t = 0 \\ ERAIN + SWAZ_{t-1} - WT_t - DR_t, t > 0 \end{cases} \quad (5)$$

where MSD is the maximum soil depth (mm), AWC is the fraction of soil water that is plant available (dimensionless), ERAIN is effective rainfall (precipitation that infiltrates into the soil on day t, mm), ET_t is the evapotranspiration (mm) on day t, DR_t is deep drainage (mm) on day t, and MSD is 300 mm deeper than RZDM, providing a soil buffer zone below the maximum RZD.

When SWAZ exceeds the water-holding capacity of the AZD (i.e., $SWAZ_t > AWC \times AZD$) the excess water is transferred to SWPZ. Likewise, when SWPZ exceeds its water-holding capacity, the excess water is lost as DR. Water dynamics in the soil layers are proportional to the relative soil water gradient (Δ_W, Equation 6) (mm), which regulates the WT between the SWPZ and SWAZ (Equations 7, 8, and 9):

$$\Delta_W = \frac{\mathrm{SWPZ}_t}{\left(\mathrm{MSD} - \mathrm{AZD}_t\right) \times \mathrm{AWC}} - \frac{\mathrm{SWAZ_t}}{\mathrm{AZD}_t \times \mathrm{AWC}} \quad (6)$$

$$\mathrm{UF}_t = \begin{cases} \Delta_W \times \mathrm{SWPZ}_t \times \mathrm{TW} \times \dfrac{\min\left(\left(\mathrm{MSD} - \mathrm{AZD}_t\right), \mathrm{AZD}_t\right)}{\max\left(\left(\mathrm{MSD} - \mathrm{AZD}_t\right), \mathrm{AZD}_t\right)}, \Delta_W \geq 0 \\ \Delta_W \times \mathrm{SWAZ}_t \times \mathrm{TW} \times \dfrac{\min\left(\left(\mathrm{MSD} - \mathrm{AZD}_t\right), \mathrm{AZD}_t\right)}{\max\left(\left(\mathrm{MSD} - \mathrm{AZD}_t\right), \mathrm{AZD}_t\right)}, \Delta_W < 0 \end{cases} \quad (7)$$

$$\mathrm{SWAZ}_t = \mathrm{SWAZ}_t + \mathrm{UF}_t \quad (8)$$

$$\mathrm{SWPZ}_t = \mathrm{SWPZ}_t - \mathrm{UF}_t \quad (9)$$

where TW is a factor of time (in days) to reach an equilibrium between the two layers, and UF_t is the unsaturated flux between the two layers (mm).

When the water supply to the plant is low, then evapotranspiration (ET) is reduced. We define a crop-specific fraction of plant-available water, dryStartET, below which the crop is drought-stressed. From this fraction, we define a soil water stress factor *DrystressET* (Equation 10), which is then used to calculate the crop's reduced ET (Equation 11).

$$\mathrm{Drystress}_{\mathrm{ET}} = \max\left(\min\left(\frac{\mathrm{SWAZ}_t \times \mathrm{dryStartET}}{\mathrm{AWC} \times \mathrm{AZD}_t}, 1\right), 0\right) \quad (10)$$

$$\mathrm{ET}_t = \min\left(\mathrm{Drystress}_{ET} \times \mathrm{K}_C \times \mathrm{ETo}_t, \mathrm{SWAZ}_t\right) \quad (11)$$

where ETo_t is the reference ET on day t (mm day^{-1}), and Kc is the ET coefficient of the current crop (dimensionless), which changes with growth stage.

Finally, excess water on the surface is calculated using the curve number method (USDA, 1986), in which a coefficient is generated to calculate the surface runoff water, that is, water that arrives at the soil surface through precipitation or irrigation but does not infiltrate.

2.1.2 | Crop growth

Like SIMPLE, the CNPP's plant growth is regulated by light, temperature, and water stress. However, the stress functions that represent reductions in biomass production due to temperature and soil available water are slightly different.

To the aboveground biomass (AGB) and yield components calculated by SIMPLE, we add computation of belowground biomass (BGB; to provide carbon input to PROCS) and a biomass decay factor for each plant compartment (aboveground and belowground) as follows:

$$\mathrm{AGB}_t = \mathrm{AGB}_{t-1} + (\mathrm{RUE} \times f\,(\mathrm{Solar}) \times \mathrm{SRAD} \times f\,(\mathrm{Temp}) \times f\left(\mathrm{CO}_2\right) \times f\,(\mathrm{Water}) \times 10) \quad (12)$$

$$\mathrm{BHARV}_t = \left(\mathrm{AGB}_t - \left(\mathrm{AGB}_{t-1} \times f(\mathrm{Decay})_c\right)\right) \times \mathrm{HI} \quad (13)$$

$$\mathrm{BLIT}_t = \left(\mathrm{AGB}_t - \left(\mathrm{AGB}_{t-1} \times f(\mathrm{Decay})_c\right)\right) \times (1 - \mathrm{HI}) \quad (14)$$

$$\mathrm{BGB}_t = \left(\mathrm{AGB}_t - \left(\mathrm{AGB}_{t-1} \times f(\mathrm{Decay})_c\right)\right) \times \mathrm{RTS} \quad (15)$$

where RUE is the radiation use efficiency (g DM MJ^{-1}), f(Solar) is the fraction of solar radiation interception (dimensionless), SRAD is the daily solar radiation (MJ m^{-2} day^{-1}), f(Temp) is a function (Equation 16) describing how RUE responds to daily air temperature (g MJ^{-1}), $f(\mathrm{CO}_2)$ is a function representing atmospheric CO_2 concentration, f(Water) is a function (Equation 17) of RUE response to soil available water status, HI is the harvest index (dimensionless ratio of harvested mass to preharvest AGB), AGB_t is the total aboveground biomass (kg DM ha^{-1}), BHARV_t is the aboveground harvested biomass (kg DM ha^{-1}), BLIT_t is the aboveground non-harvested biomass (added to litter after harvest events, kg DM ha^{-1}), BGB_t is the belowground biomass (kg DM ha^{-1}), RTS is the fixed root to shoot factor that calculates root biomass in relation to AGB (dimensionless), and $f(\mathrm{Decay})_c$ is a mortality factor for each plant compartment (subscript "c").

The f(Temp) considered in CNPP is a single trapezoidal function, whereas two separate temperature equations are used in SIMPLE. The f(Temp) function (Equation 16) is a trapezoidal function that adjusts RUE given the daily mean air temperature T_{mean} and four reference temperatures (°C): T_{base} is the basal temperature below which daily biomass production is zero, T_{opt1} and T_{opt2} are, respectively, the lower and upper ends of the optimal temperature range within which daily biomass production is maximized, and T_{ext} is the extreme (high) temperature above which daily biomass production is zero.

TABLE 1 Summary of the datasets used for the parameterization, calibration, and evaluation of the carbon net primary productivity (CNPP) module.

Dataset	Crop	Coordinates	Köppen climate classification	Level of information
1	Soybean Maize	41.1647 N, 96.47 W (US-Ne2) and 41.1794 N, 96.4394 W (US-Ne3)	Dfa	LAI, ET, biomass
2	Wheat	27.9061–30.3408 S, 51.0214–54.8881 W	Cfa, Cfb	Yield
3	Common bean	9.4672–28.5092 S, 35.8286–55.80 W	As, Aw, Cwa, Cfa, Cwb, Cfb	Yield
4	Perennial forage	22.70 S, 47.50 W	Cwa	Biomass
5	Soybean	3.0636–33.1964 S, 44.1225–62.9614W	Af, Am, Aw, Cfa, Cfb, Cwa, Cwb	Yield
6	Maize	1.7775–33.1978 S, 36.8019–60.8611 W	Af, Am, As, Aw, Cwa, Cwb, Cfa, Cfb	Yield
7	Wheat	12.1439–30.3392 S, 43.7703–55.7950 W	Cwa, Cfa, Cwb, Cfb, Aw	Yield
8	Common bean	7.50–29.20 S, 43.7597–57.4131 W	Am, Aw, Cwa, Cfa, Cwb, Cfa	Yield

Abbreviations: ET, evapotranspiration; LAI, leaf area index.

$$f(\text{Temp}) = \begin{cases} \max\left(\frac{T_{\text{mean}} - T_{\text{base}}}{T_{\text{opt1}} - T_{\text{base}}}, 0\right), T_{\text{mean}} < T_{\text{opt1}} \\ 1, T_{\text{opt1}} < T_{\text{mean}} < T_{\text{opt2}} \\ \max\left(1 - \frac{T_{\text{mean}} - T_{\text{opt2}}}{T_{\text{ext}} - T_{\text{opt2}}}, 0\right), T_{\text{mean}} > T_{\text{opt2}} \end{cases} \quad (16)$$

The *f*(Water) is a dry stress factor determining the soil water status in relation to that fraction of plant-available water at which stress starts to reduce plant growth (Equation 17).

$$f(\text{Water}) = \min\left(\frac{\text{SWAZ} \times \text{dryStartGR}}{(\text{DMX} - \text{AZD}_t) \times \text{AWC}}, 1\right) \quad (17)$$

where dryStartGR is the fraction of plant-available water below which drought stress reduces plant growth (dimensionless).

2.2 | Field dataset descriptions

CNPP calibration and evaluation were performed for four annual crops and one perennial forage species: soybean [*Glycine max* (L.) Merr.], maize (*Zea mays L.*), wheat (*Triticum aestivum* L.), common bean (*Phaseolus vulgaris* L.), and perennial forage [*Urochloa* (syn. *Brachiaria*) *brizantha* (Hochst ex A. Rich.) Stapf cv. Marandu]. Data for modeling and model evaluation for soybean and maize were based on data collected in a well-monitored micrometeorological experiment (dataset 1). Wheat, bean, and "Marandu" palisade grass input data were collected in agrometeorological experiments, respectively datasets 2, 3, and 4 (Table 1). Additional datasets from a breeding program, representing on-farm conditions (datasets 5, 6, 7, and 8, respectively, for soybean, maize, wheat, and bean), were distributed across the most important Brazilian agricultural areas (Figure 1). For these datasets, only grain yield data were available; no information was given for other biomass components such as aboveground residue.

Dataset 1 is part of a long-term field experiment conducted at the Agricultural Research and Development Center of the University of Nebraska-Lincoln near Mead, NE, USA, which has long-term hourly measurements of surface fluxes obtained via eddy covariance, soil moisture monitoring, crop phenology, and biomass time series observations. Two sites were used: (1) US-Ne2 (41.1647 N, 96.47 W, elevation 362 m, mean average temperature 10.08°C, and mean annual precipitation 788.89 mm) and (2) US-Ne3 (41.1794 N, 96.4394 W, elevation 363 m, mean average temperature 10.11°C, and mean annual precipitation 783.68 mm). The US-Ne2 site is irrigated with a center pivot system, while US-Ne3 relies on rainfall (Suyker, 2024a, 2024b). The climate is designated Dfa (humid continental: humid with severe winter, no dry season, hot summer) in the Köppen climate classification system. The soils are deep silty clay loams (Mollic Hapludalfs, Pachic Argialbolls, and Vertic Argialbolls). Data from both sites (US-Ne2 and US-Ne3) are based on a maize–soybean crop rotation system under no-till since 2001 (details at https://ameriflux.lbl.gov/). For the present study, data from 2001 to 2014 were used, varying the crop and the irrigation scheme.

Dataset 2 is from wheat breeding trials and includes information on grain yield, cultivar cycle, planting and harvesting dates, and experimental location. This dataset includes nonirrigated experiments conducted during the off-season (winter season) between 2009 and 2022 at eight sites (Figure 1): five in Brazil's wheat producing region I (WPR I) (Cunha et al., 2006), and three in WPR II. A total of 132 cultivars, ranging from early to medium growth cycles, were evaluated in these experiments. Thus, dataset 2 comprises a total of 4745 experiments: the calibration subset (totaling 2948 experiments) includes data from three sites in WPR I and two in WPR II,

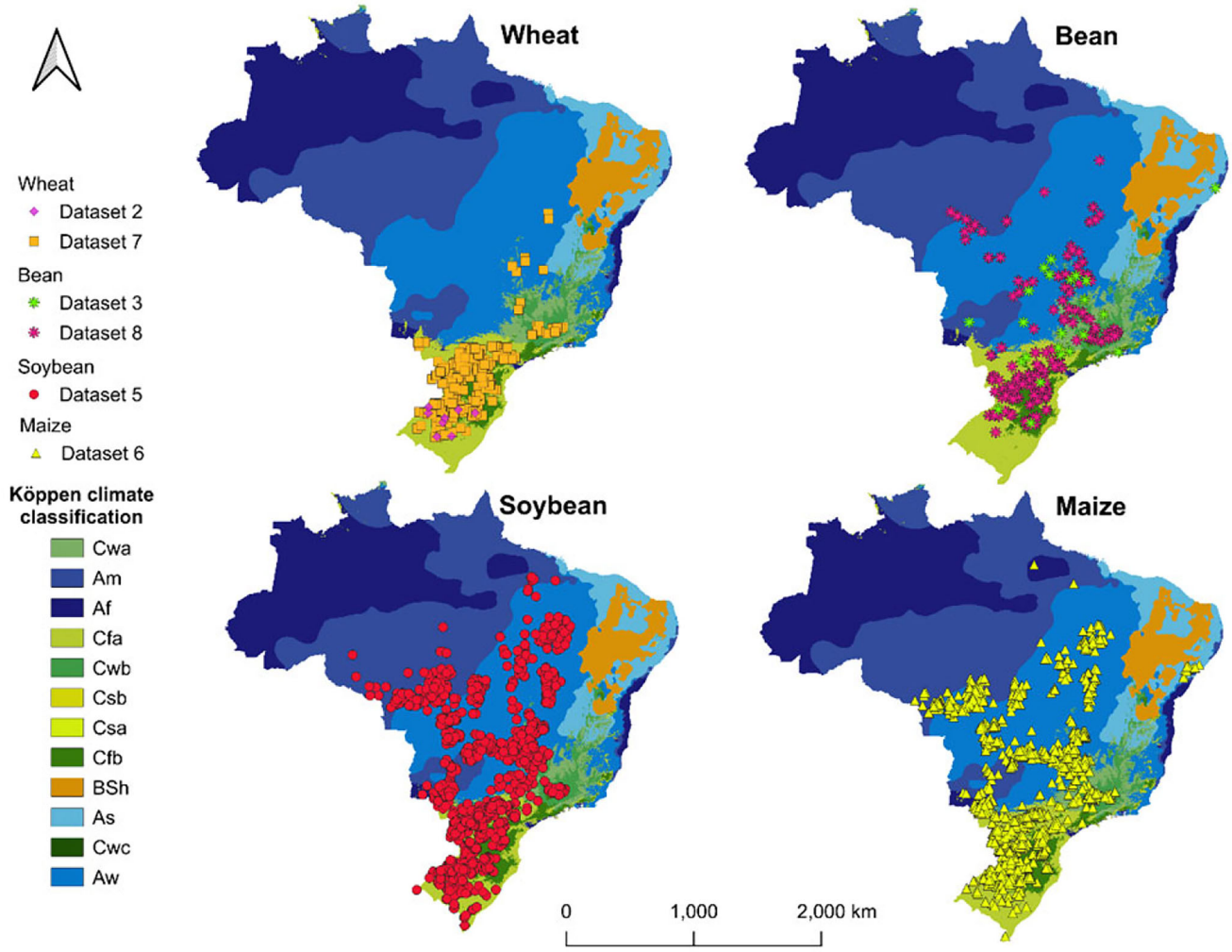

FIGURE 1 Agrometeorological sites and on-farm plots used for parametrizing and evaluating the carbon net primary productivity (CNPP) in Brazil. Köppen climate classification map is based on Alvares et al. (2013). Datasets 1 and 4 consist of individual data points situated in the Northern and Southern Hemispheres, respectively, and are consequently omitted from the map visualization.

while the evaluation subset (1797 experiments) contains data from two sites in WPR I and one in WPR II.

Dataset 3 contains data on common bean obtained from the literature in Table S2 from irrigated and nonirrigated experiments conducted at various times between 2000 and 2018. This dataset contains information on grain yield, cultivar cycle, planting and harvesting dates, and experimental location. Dataset 3 includes a total of 71 cultivars from 23 sites (Brazilian municipalities; Figure 1), covering growth habits I, II, and III. This dataset comprised a total of 161 experiments and was used only for module calibration.

Dataset 4 comprised data on the perennial forage collected from an experimental area at the University of São Paulo, Luiz de Queiroz College of Agriculture in Piracicaba, São Paulo, Brazil (22°84′20″ S, 47°83′00″ W). The experiment was conducted between April 2011 and April 2013 under four treatments, the product of (1) irrigated versus nonirrigated and (b) mechanized biomass harvesting every ~28 days versus every ~42 days. Further details on the experiment, data collection, and irrigation management can be found in Pequeno et al. (2014) and Pequeno et al. (2018). The main collected variables included herbage dry mass production (biomass harvested above a 10-cm stubble height), live perennial forage dry mass remaining (stubble dry mass), leaf area index (LAI), canopy light interception (LI), and harvest date. Weather data were obtained from a weather station located 1.8 km from the experimental area.

For soybean (dataset 5; breeding trial data), a split-block design was used, having 50,900 crop records (unique combination of genotype, cultivation year, and site location) across 509 sites (municipalities), nested within 32 micro-regions. Data were collected from 2017 to 2024 across two growing seasons: rainy season, the most productive period; and off-season, a dry and less productive period. The dataset has 572 coded genotypes binned into five relative maturity (RM) groups, designated RM5 through RM9. All fields have

information on yield, and planting and harvest dates, with a subset of 14,373 records also having information on date of physiological maturity. This subset was used in the calibration procedure to precisely estimate soybean lifespan and harvest date. Further information is available in Section 2.3.

Dataset 6 is structured similarly to dataset 5: its 24,139 maize records (unique combinations of genotype, cultivation year, and site location), located in 433 sites, are binned into five mega environments (or breeding regions): "Subtropical Late," "Subtropical Summer," "Subtropical Winter," "Tropical Highlands," and "Tropical Lowlands." Data were collected from 2017 to 2024 across rainy season and offseason. A subset of 4443 records also has information on silking date (the developmental stage when silks are first visible), which was used to estimate physiological maturity and harvest date. Further information is available in Section 2.3.

Datasets 7 (260 wheat sites) and 8 (153 common bean sites) were collected from a breeding program conducted on rainfed farm fields during main- and offseasons between 2020 and 2023. The wheat sites were distributed across WPR I, II, III, and IV, and include cultivars ranging from early to medium growth cycles. Common bean sites were in the main Brazilian bean producing regions (Figure 1); however, information on cultivars and growth habits was not available. Datasets 7 and 8, respectively, comprise 3409 and 1071 crop records.

2.3 | Parametrization and evaluation procedures

For soybean and maize, dataset 1 has the most detailed measurements ("full parametrization"), even including a flux tower, so it was used to parametrize the biophysical and physiological components of the module. Module parametrization was performed using a hierarchical structure of simulations wherein a combination of parameters was tested varying in predefined intervals obtained from Zhao et al. (2019). A hierarchical sequence was followed, proceeding from calibration of biophysical processes (radiation interception and absorption), through physiological (ET and net primary production) and phenological (TT calibration) parameters, ending with carbon allocation to plant components. The parameter set was calibrated using data exclusively from the irrigated experiments, where water stress was minimized or controlled. Parameters were then subjected to independent evaluation and validation using data from the rainfed experiments, which introduced variable water stress conditions.

The initial full parametrization aimed to refine the core biophysical, physiological, and phenological processes within the CNPP module. Building upon this established general parameter set, specific calibrations were subsequently performed at the cultivar level, for groups aggregated by maize and soybean production regions. Notably, the specific calibration for soybean relied exclusively on dataset 5, where specific parametrizations were performed using combinations of RM and micro-regions. Parameters regulating (1) plant lifespan (cumulative temperature requirement from planting to maturity), (2) cumulative temperature requirement that regulates the plant's light interception to reach 50% of maximum radiation interception during the early growth phase, (3) cumulative temperature requirement that regulates the plant's light interception to reach 50% of maximum radiation interception during leaf senescence, (4) maximum LI, (5) root to shoot, and (6) HI were all changed during this specific calibration (Table 2). The adjusted parameters for these parametrization are available in Table S3. This dataset was partitioned into mutually exclusive calibration and evaluation subsets. The calibration set was used to develop and refine the approach, while the evaluation set was used for independent evaluation of its performance, avoiding data cross-contamination that would compromise validation. Temporal variability was explicitly incorporated into the datasets by including site observations spanning the period from 2017 to 2024. This longitudinal data inclusion should minimize potential biases associated with overly homogeneous climatic conditions.

Similar calibration/evaluation procedures were performed for maize using dataset 6, where a specific parametrization was performed for each MG group in the mega-environments. The full parametrization was used as reference for those parameters that regulate plant lifespan; parameters that were recalibrated regulated maximum LI, water stress, and HI.

For wheat, common bean, and perennial forage, an initial parametrization was conducted based on the literature (Table S2). Subsequently, module calibration and evaluation were performed starting from the initial parametrization and working from the intermediate-quality datasets 2, 3, and 4.

Module calibration for wheat was carried out separately for each combination of cycle classification (early and medium growth cycles) and WPR (I and II), using the dataset 2. Model performance was then evaluated using dataset 7. For WPRs III and IV, which were not represented in dataset 2, a separate calibration was conducted using the calibration subset of dataset 7. This recalibration step adjusted parameters governing plant lifespan, water stress, maximum LI, and HI. Finally, the model for these regions was validated using the corresponding evaluation set from dataset 7.

For common bean, the full dataset 3 was used exclusively for calibration. This calibration was conducted for each growth habit (I, II, and III), and involved regulating the same parameters used in the soybean and wheat. Subsequently, a recalibration was performed, where dataset 8 (breeding program) was split into calibration (548 crop records and 76 sites) and evaluation (523 crop records and 77 sites) subsets.

TABLE 2 Summary of the crops and datasets used for the calibration and evaluation of the carbon net primary production (CNPP) module.

	Crops/datasets				
	Wheat	Bean	Perennial forage	Soybean	Maize
Parameter	Datasets 2 and 7	Datasets 3 and 8	Dataset 4	Dataset 5	Dataset 6
RZDM (mm)	+	+	–	–	–
T_{base} (°C)	+	–	–	–	–
T_{opt1} (°C)	+	+	+	–	–
T_{opt2} (°C)	+	+	+	–	–
ExtremeT (°C)	+	+	+	–	–
RUE (g MJ^{-1})	+	+	+	–	–
Tsum (°C $period^{-1}$)	+	+	+	+	+
TbaseD (°C $cicle^{-1}$)	+	+	+	–	–
I50A (°C $period^{-1}$)	+	+	–	+	+
I50B (°C $period^{-1}$)	+	+	–	+	+
I50maxW (°C day)	+	+	–	–	–
MaxIntercept	+	+	+	+	+
DryStartET	+	+	+	–	–
DryStartGR	+	+	+	–	–
Kcmin	+	+	+	–	–
Kcmax	+	+	+	–	–
RTS	+	+	–	–	+
HI	+	+	–	+	+

Note: "+" changes applied identically to both calibration and evaluation datasets.
Abbreviations: HI, harvest index; RTS, root to shoot factor; RUE, radiation use efficiency; RZDM, maximum root zone depth.

Similarly to wheat, the recalibrated parameters regulated plant lifespan, water stress, maximum LI, and HI.

Finally, for perennial forage, local module calibration was conducted using the dataset 4. Calibration was performed using the irrigated dataset, and rainfed experiments were used for evaluation. This calibration involved adjusting the parameters related to (1) plant lifespan (cumulative temperature requirement from planting to maturity), (2) cumulative temperature requirement that regulates the plant's light interception to reach 50%, (3) maximum LI, (4) water stress, and (5) optimal temperature for biomass production. To adapt the module to a perennial crop—starting a new cycle after each biomass harvest, rather than starting from planting as with grain crops—a virtual planting date was set; this date was estimated to make the modeled *f*(Solar) be as close as possible to the *f(Solar)* estimated from the biomass observed immediately postharvest.

2.4 | Module evaluation

We use the linear modeling capability of *R* (function "lm"; R Core Team, 2024) to evaluate the module's performance in predicting crop yields. For each calibration or validation instance, we calculated the coefficient of determination (R^2; Equation 18), root mean square error (RMSE; Equation 21), relative root mean square error (RRMSE; Equation 19), mean bias (MB; Equation 20), and Nash–Sutcliffe Efficiency (NSE; Equation 21):

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (a_i - \hat{a}_i)^2}{\sum_{i=1}^{n} (a_i - \bar{a}_i)^2} \tag{18}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{a}_i - a_i\right)^2}{n}} \tag{19}$$

$$\text{RRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{n} \left(\hat{a}_i - a_i\right)^2}{n}}}{\frac{\sum_{i=1}^{n} \hat{a}_i}{n}} \times 100 \tag{20}$$

$$\text{MB} = \sum_{i=1}^{n} \frac{\left(\hat{a}_i - a_i\right)}{n} \tag{21}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n} (a_i - \hat{a}_i)^2}{\sum_{i=1}^{n} (a_i - \bar{a}_i)^2} \tag{22}$$

TABLE 3 Statistics (mean, with standard deviation in parentheses) of carbon net primary productivity (CNPP) simulations for leaf area index (LAI; $m^2\ m^{-2}$), evapotranspiration (ET; mm day^{-1}), and aboveground biomass (kg ha^{-1}) for maize and soybean in dataset 1. "*n*" is the number of harvest years.

Crop	Site	Variable	RMSE	RRMSE (%)	MB	NSE	R^2	*n*	Figure
Soybean	Calibration	LAI ($m^2\ m^{-2}$)	0.1 (0.03)	5.4 (0.5)	0.01 (0.01)	0.88 (0.1)	0.9 (0.1)	5	2
		ET (mm day^{-1})	0.8 (0.1)	32.9 (3.8)	0.2 (0.1)	0.75 (0.1)	0.8 (0.03)	5	3
		Biomass (kg ha^{-1})	141.6 (69.3)	3.6 (1.6)	12.7 (14.5)	0.95 (0)	0.97 (0.02)	5	4
	Evaluation	LAI ($m^2\ m^{-2}$)	0.2 (0.1)	7.7 (2.9)	0.02 (0.01)	0.56 (0.4)	0.9 (0.05)	7	2
		ET (mm day^{-1})	0.9 (0.2)	37.5 (10.8)	0.2 (0.1)	0.67 (0.2)	0.76 (0.1)	7	3
		Biomass (kg ha^{-1})	190.7 (74.4)	6 (2.2)	13 (26.4)	0.85 (0.1)	0.95 (0.04)	7	4
Maize	Calibration	LAI ($m^2\ m^{-2}$)	0.1 (0.05)	4.3 (1.1)	0.0002 (0.02)	0.89 (0.1)	0.94 (0.05)	9	5
		ET (mm day^{-1})	0.8 (0.1)	32.3 (5.1)	0.1 (0.2)	0.79 (0.1)	0.81 (0.04)	9	6
		Biomass (kg ha^{-1})	297.3 (112.6)	3 (1.3)	30.6 (40.2)	0.97 (0.02)	0.99 (0.01)	9	7
	Evaluation	LAI ($m^2\ m^{-2}$)	0.1 (0.05)	4.6 (1.5)	0.01 (0.01)	0.85 (0.1)	0.92 (0.05)	6	5
		ET (mm day^{-1})	0.8 (0.1)	38 (4.5)	0.1 (0.1)	0.72 (0.1)	0.78 (0.05)	6	6
		Biomass (kg ha^{-1})	333 (84.5)	4.1 (1.2)	23.1 (35.9)	0.93 (0.02)	0.98 (0.01)	6	7

Abbreviations: MB, mean bias; NSE, Nash–Sutcliffe efficiency; R^2: coefficient of determination; RMSE, root mean square error; RRMSE, relative root mean square error.

where a_i denotes the observed values, $\hat{a}_i$ the predicted values, $\bar{a}$ the average of the observed values, and n the number of observations.

3 | RESULTS

3.1 | Full CNPP parametrization (soybean and maize)

High performance, characterized by high R^2 values and low RMSE, RRMSE, and MB values (Table 3), was achieved in the simulations of LAI, ET, and AGB for soybean (Figures 2, 3, and 4) and corn (Figures 5, 6, and 7) in dataset 1. Consistency between the calibration dataset (US-Ne2) and the evaluation dataset (US-Ne3) was evident for both crops, suggesting that the parameters derived from the calibration conducted on the irrigated site (US-Ne2 used for calibration) effectively predicted biomass on the rainfed site (US-Ne3 used for evaluation) (Figures 4 and 7). We emphasize that all analyses of LAI, ET, and biomass presented in Table 3 were based on the whole lifespan.

3.2 | CNPP parametrization for low-data crops (wheat, bean, and perennial forage)

The strong performance of the CNPP module seen in Section 3.1, when the module was parametrized using dataset 1 (complete local values including flux tower data), is here contrasted with lower performance when parametrization is based on datasets with limited information content. Specifically, parametrization for wheat, common bean, and perennial forage was partially based on literature-derived parameter values. Moreover, the wheat (dataset 2) and common bean (dataset 3) module calibrations lacked long-term measurements and were limited to yield observations only.

A further source of uncertainty arose because these field experiments were breeding trials that seldom included concurrent local climate and soil data collection. Consequently, the relationship between simulated and observed wheat and bean yields (Figure 8) exhibits greater variability, as crucial variables such as LAI and ET could not be rigorously evaluated.

A lower predictive performance was noted for wheat (Table 4). Furthermore, the module exhibited low predictive accuracy for observed biomass yield of the perennial forage (Figure 8), across both irrigated (calibration) and rainfed (evaluation) conditions.

3.3 | Evaluating regional-scale yield simulations for Brazil

The modal cultivar calibrations—which utilized a single parameter set representing all cultivars within a given crop (soybean, maize, wheat, and common bean)—were conducted using a regionalized approach, that is, combining all locations within a region. Consequently, the primary evaluation of module performance was conducted at this regional scale. Detailed analysis of the module's performance at the local scale (comparing simulations against observed yield at individual sites) for all crops is available in Table S4 and Figure S2. The regionalized results (Table 5) therefore represent mean yields derived from aggregating measurements

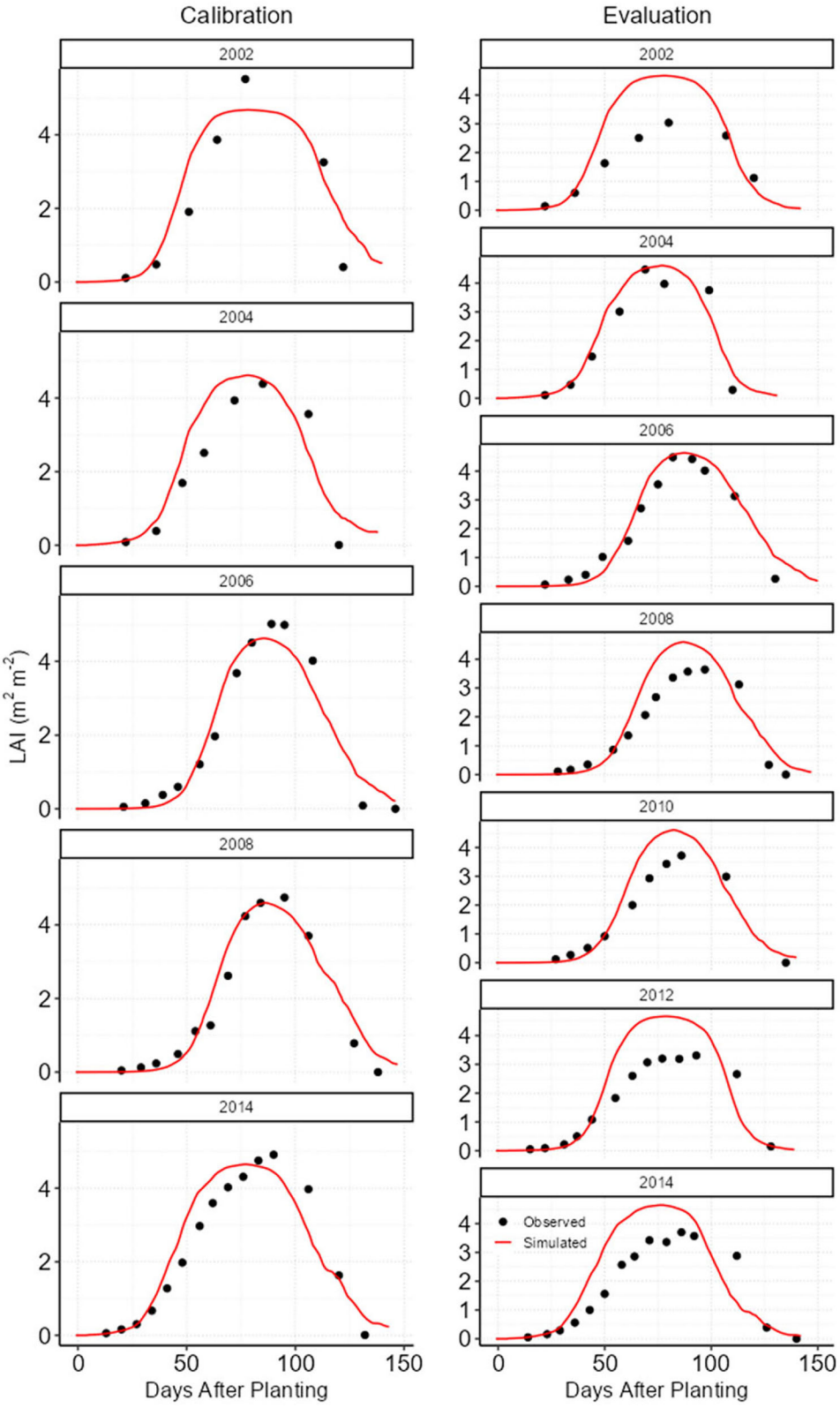

FIGURE 2 Simulated and observed soybean leaf area index (LAI) at the dataset 1 sites used for calibration (left) and evaluation (right).

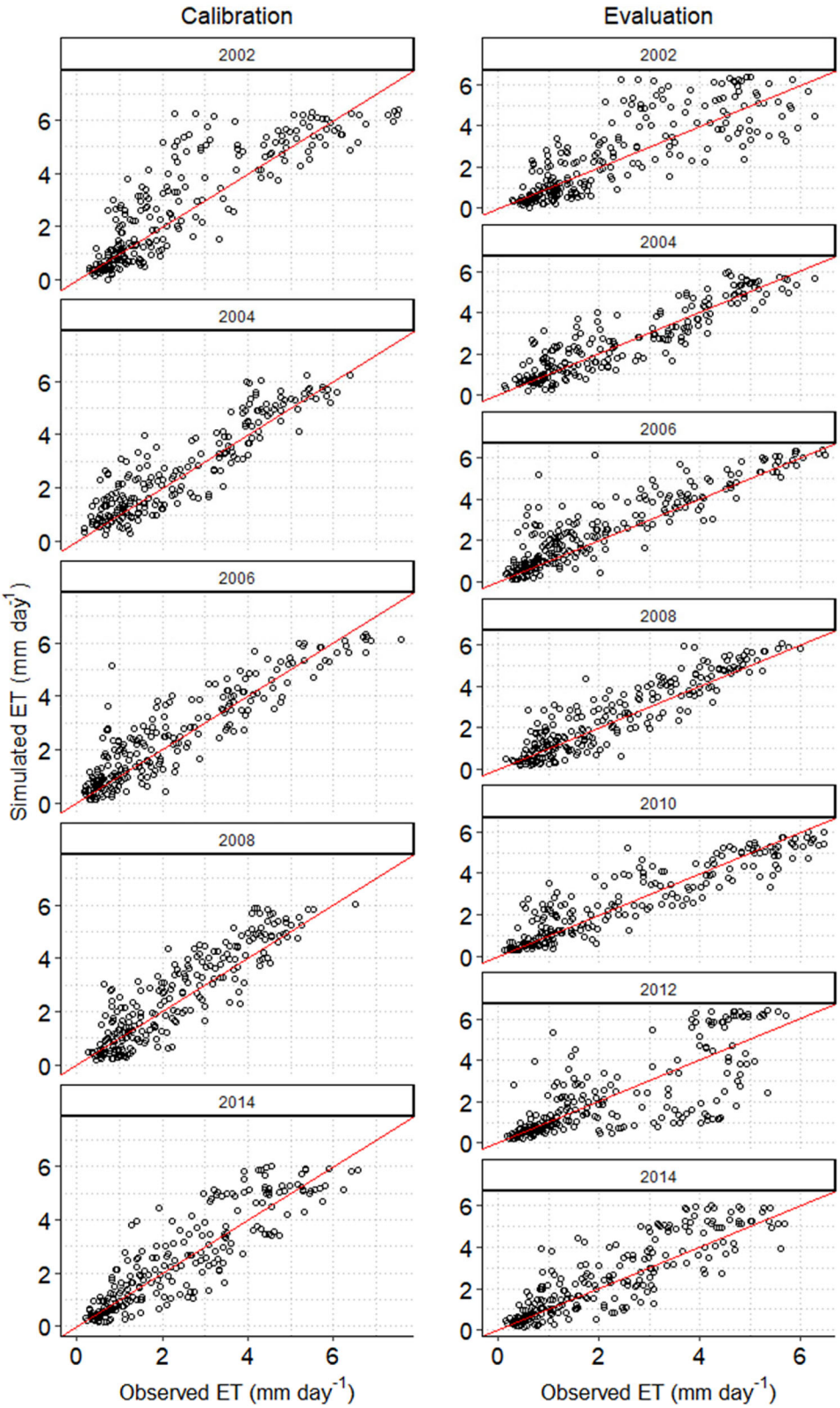

FIGURE 3 Simulated and observed soybean evapotranspiration (ET) at the dataset 1 sites used for calibration (left) and evaluation (right). Red line is the 1:1 line.

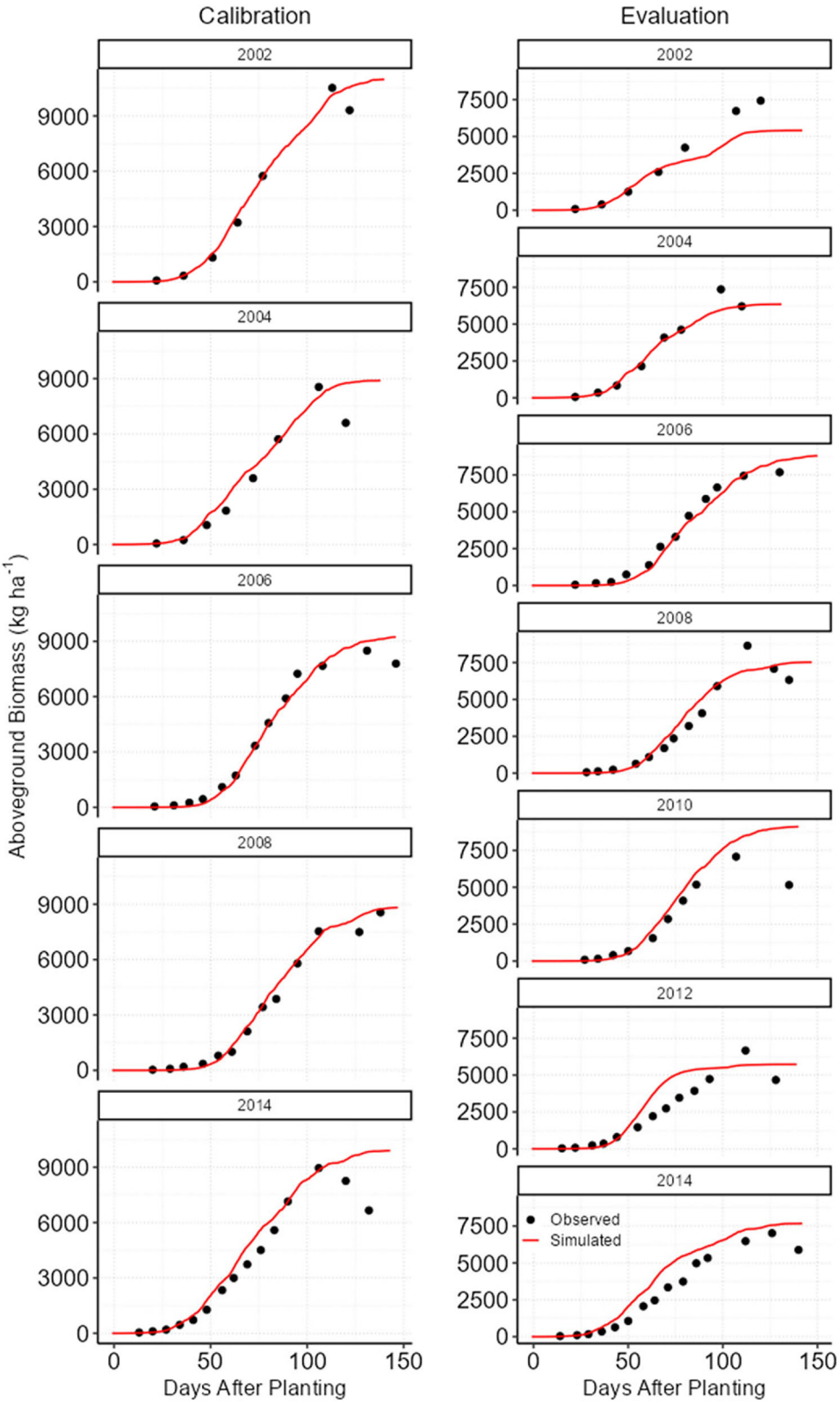

FIGURE 4 Simulated and observed soybean aboveground biomass at the dataset 1 sites used for calibration (left) and evaluation (right).

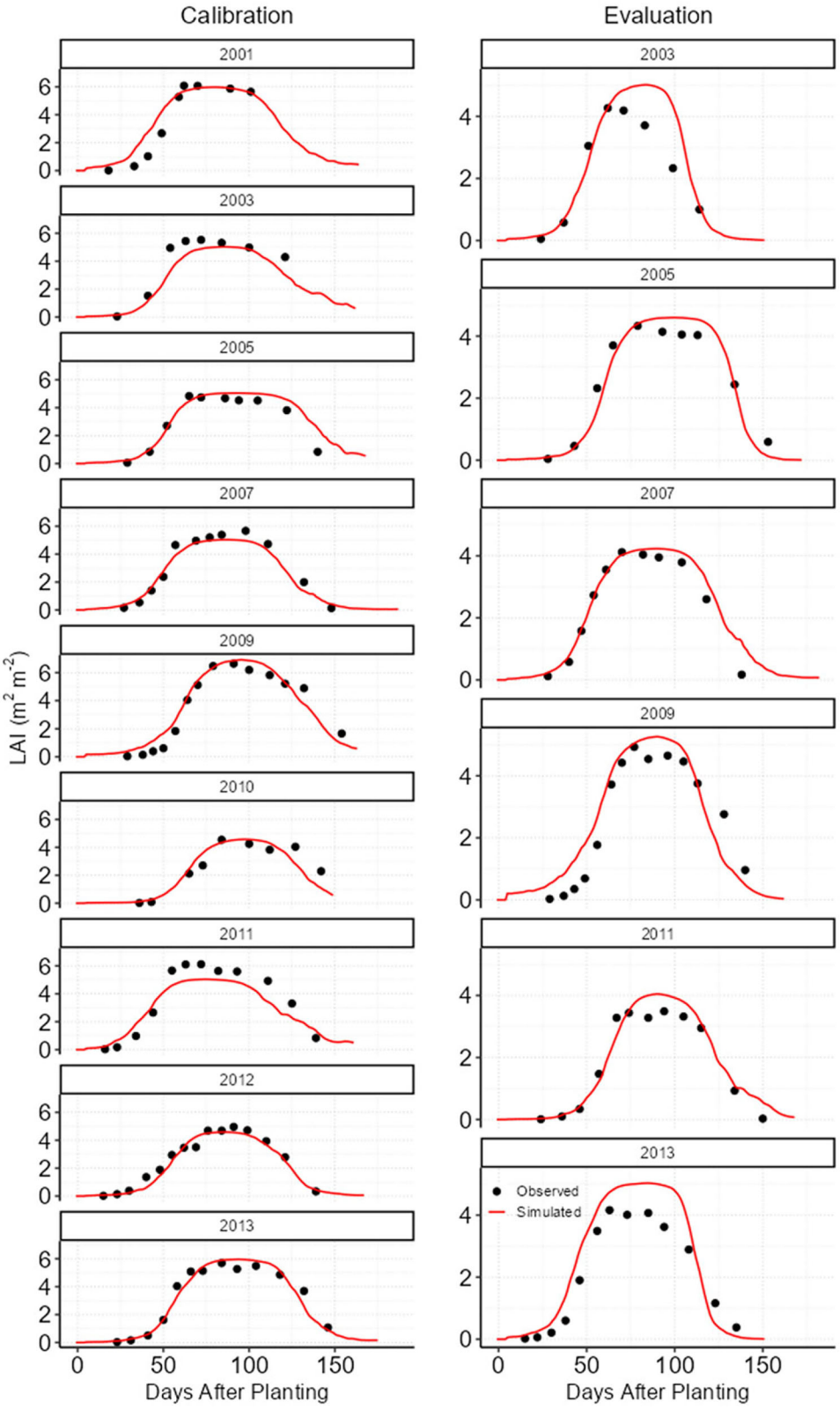

FIGURE 5 Simulated and observed maize leaf area index (LAI) at the dataset 1 sites used for calibration (left) and evaluation (right).

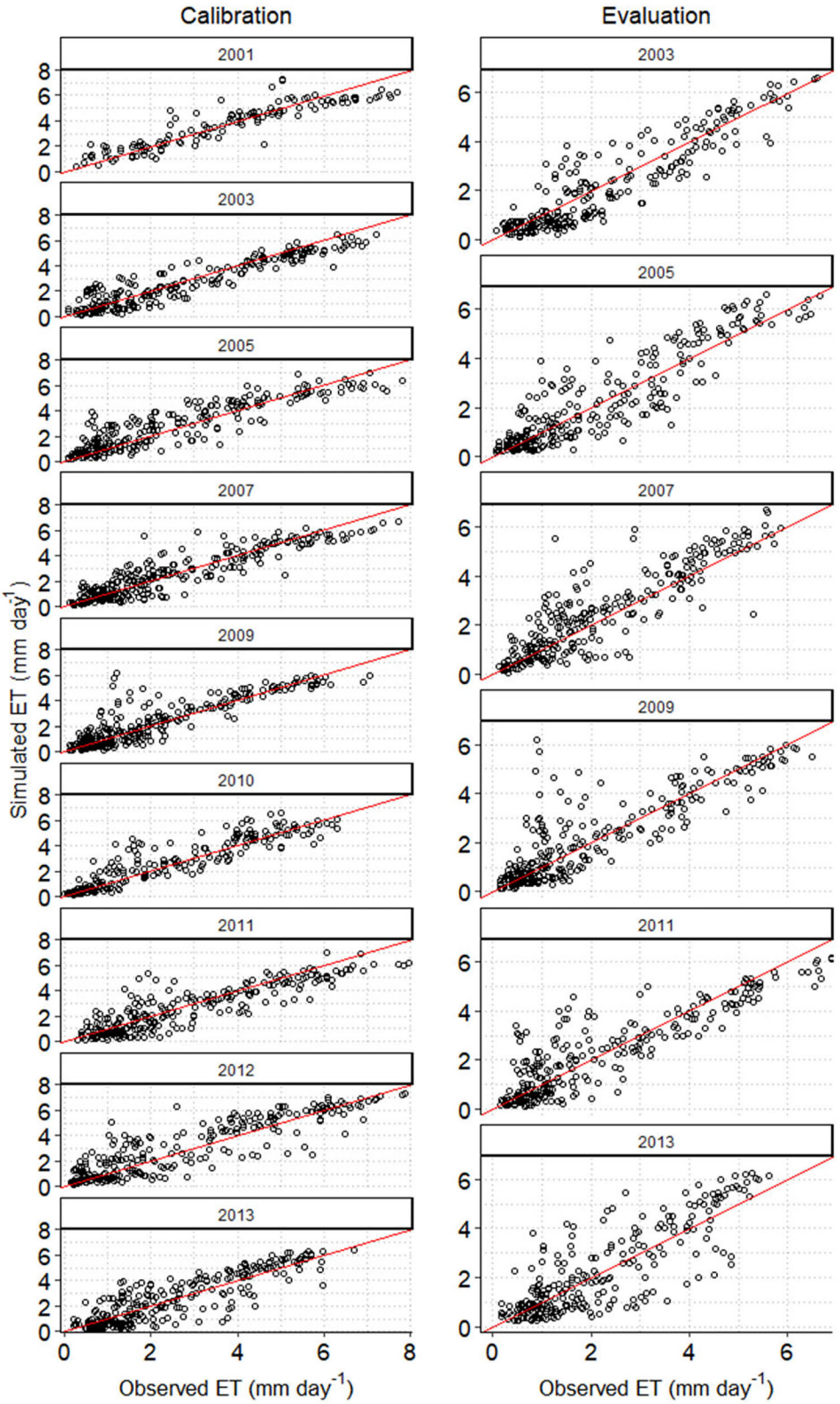

FIGURE 6 Simulated and observed maize evapotranspiration (ET) at the dataset 1 sites used for calibration (left) and evaluation (right). Red line is the 1:1 line.

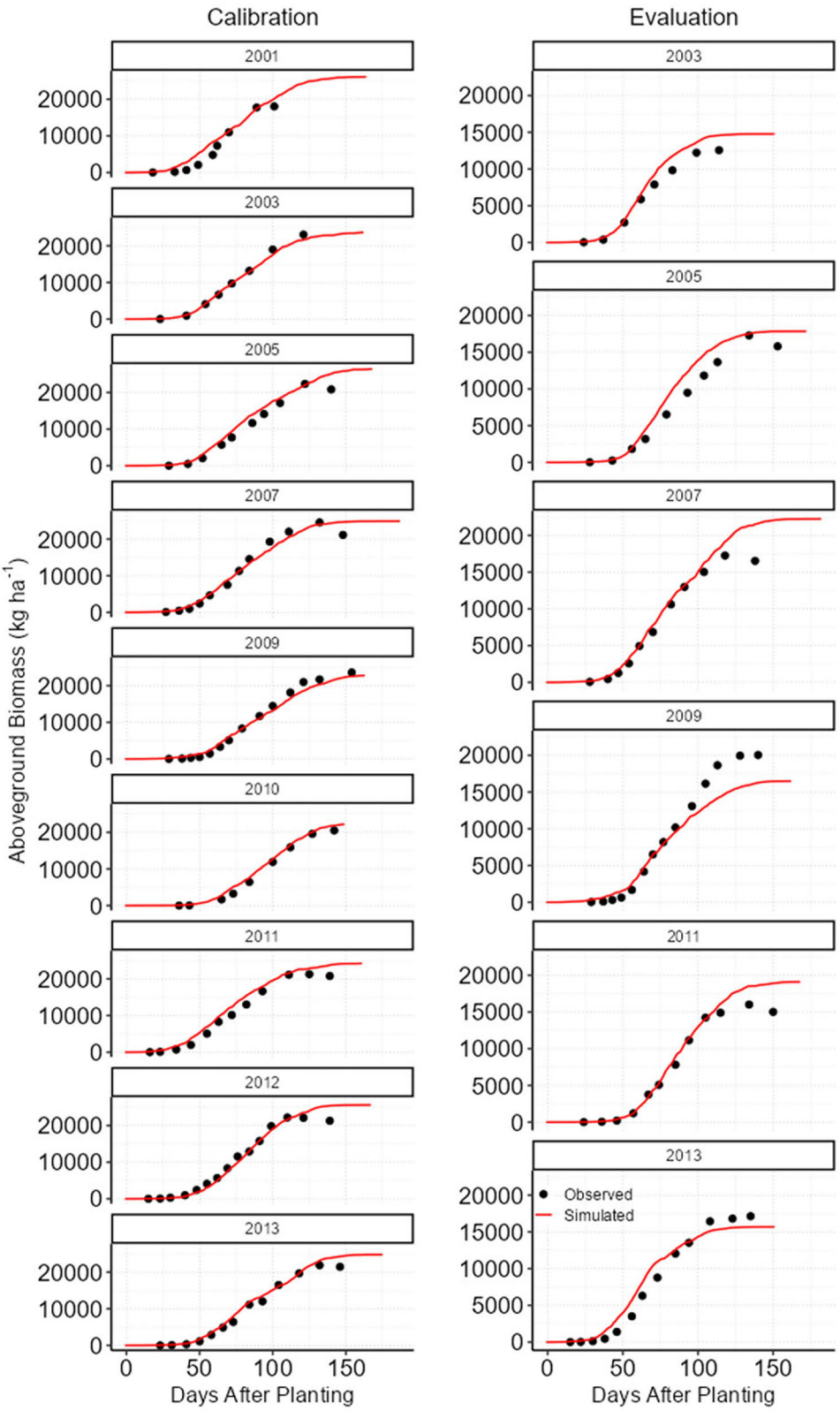

FIGURE 7 Simulated and observed maize aboveground biomass at the dataset 1 sites used for calibration (left) and evaluation (right).

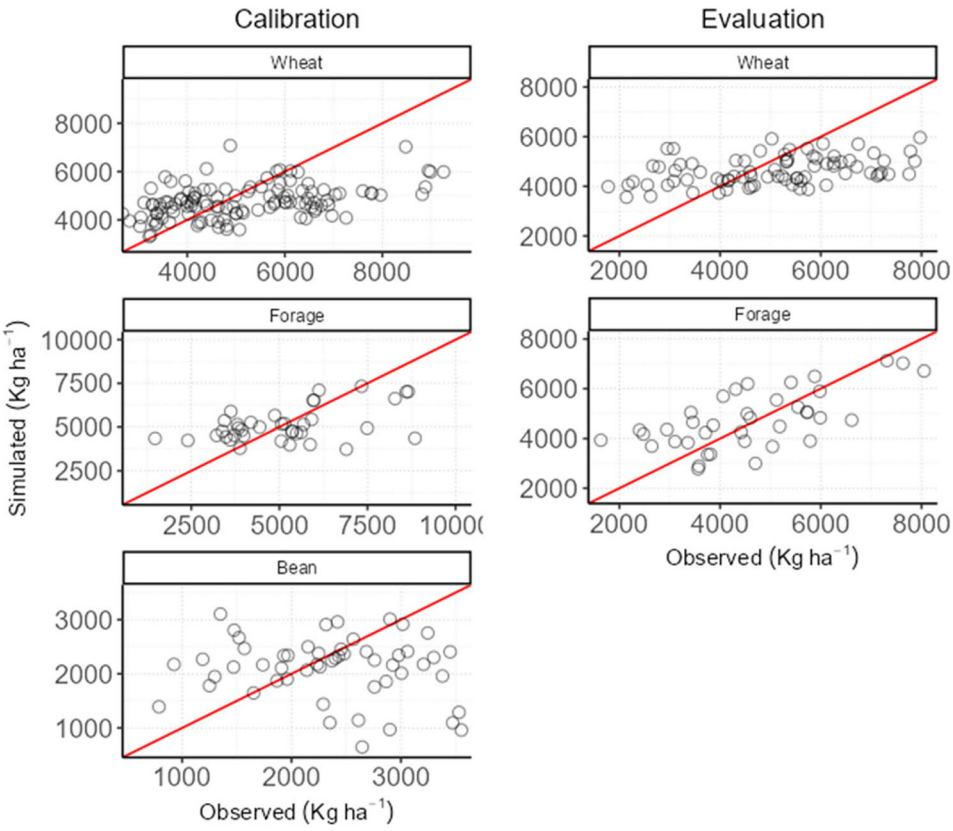

FIGURE 8 Simulated and observed grain yield of wheat and common bean, and simulated and observed biomass yield for perennial forage. Red line is the 1:1 line.

TABLE 4 Statistics of carbon net primary productivity (CNPP) simulations of yield (kg ha^{-1}) for wheat and common bean, and biomass yield for perennial forage in experimental sites (datasets 2, 3 and 4, respectively).

Crop	Subset	RMSE	RRMSE (%)	MB	NSE	R^2
Wheat	Calibration	1453.4	30.5	−382.1	0.14	0.2
	Evaluation	1474.3	31.8	−399.7	0.10	0.17
Forage	Calibration	1514.2	29.6	−19.6	0.24	0.25
	Evaluation	1135.3	24	145	0.39	0.42
Bean	Calibration	1029.4	48.5	−318.7	−0.81	0.01

Abbreviations: MB, mean bias; NSE, Nash–Sutcliffe efficiency; R^2: coefficient of determination; RMSE, root mean square error; RRMSE, relative root mean square error.

TABLE 5 Statistics of carbon net primary productivity (CNPP) regionalized yield simulations (Mg ha^{-1}) for maize, soybean, wheat, and common bean from datasets 5, 6, 7 and 8, respectively.

Crop	Subset	RMSE	RRMSE	MB	NSE	R^2
Soybean	Calibration	0.59	18.37	0.07	0.23	0.24
	Evaluation	0.64	19.76	0.09	0.02	0.09
Maize	Calibration	0.78	11.31	0.03	0.72	0.72
	Evaluation	0.94	13.7	0.04	0.64	0.65
Wheat	Calibration	0.77	25.85	0.28	−0.12	0.29
	Evaluation	0.87	30.59	−0.22	−0.82	0.19
Bean	Calibration	0.47	22.49	0.1	−2.72	0.2
	Evaluation	0.65	30.24	0.28	−12.0	0.06

Abbreviations: MB, mean bias; NSE, Nash–Sutcliffe efficiency; R^2: coefficient of determination; RMSE, root mean square error; RRMSE, relative root mean square error.

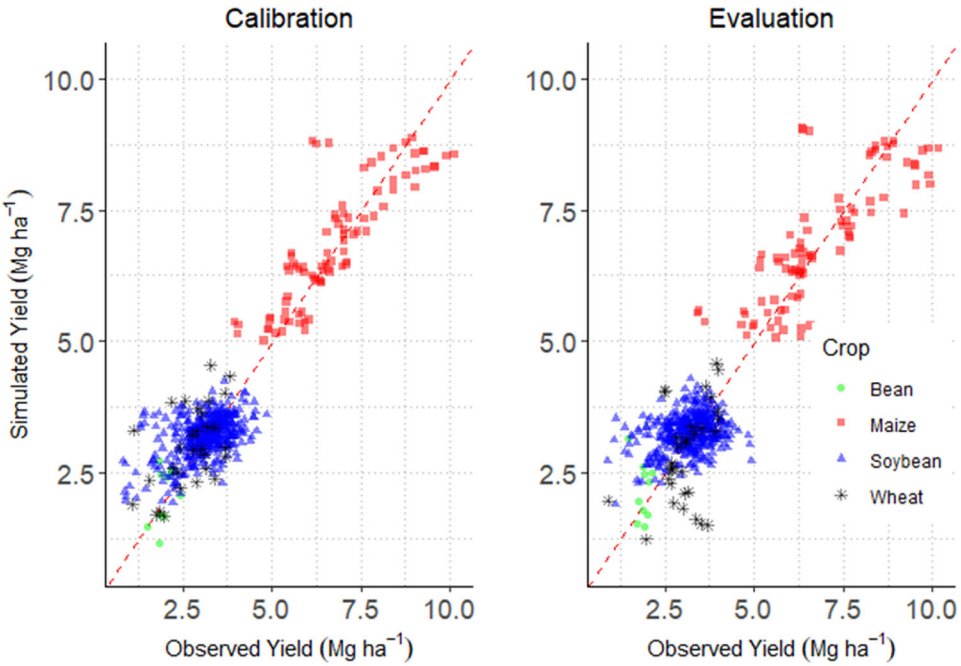

FIGURE 9 Simulated and observed regionalized yield (Mg ha^{-1}) for maize, soybean, wheat, and common bean from datasets 5, 6, 7, and 8 respectively. Red line is the 1:1 line.

within subsets of genotype (GMR for soybean and MG for maize and wheat) and regions (micro-regions for soybean and mega-environments for maize), across crop seasons. Figure 9 illustrates these groupings, specifically showing mean yields for soybean, maize, wheat, and common bean.

As observed, regional-scale simulations (Table 5; Figure 9) demonstrated improved performance mostly for maize, with intermediate performance for soybean and wheat, and a lower performance for bean.

4 | DISCUSSION

The CNPP module was initially calibrated for soybean and maize using a high-quality dataset, with flux tower measurements and in-season biomass data. The module shows consistent accuracy for simulating ET, LAI, and AGB. Comparisons with published studies revealed that CNPP's performance was either equivalent to or better than several models in the literature for soybean (Silva et al., 2025) and maize (Amiri et al., 2022; Kimball et al., 2023) at the same or similar experimental sites. Robust performance was observed for both calibration (irrigated) and validation (rainfed) datasets, thereby confirming the effectiveness and robustness of the newly implemented modifications, especially the water stress functions.

In contrast, to the CNPP's high performance for calibrating and evaluating the full parametrization (dataset 1), the calibration to Brazilian conditions (datasets 5 and 6) resulted in a decreased performance. The calibration process performed using on-farm data (datasets 5 and 6) largely had only yield data, with phenological information available only for a small subset. For these specific calibrations, phenology served as the basis for genotype clustering, that is, genotypes exhibiting similar phenologies within similar regions were clustered together. The parameters tuned during this process primarily targeted plant lifespan, maturity, and leaf senescence, which are known to strongly influence biomass development and allocation to various plant compartments. We suggest that plant development as a function of phenology is not uniform across all genotypes within a given cluster (except for maize), even if their overall phenological timing is similar. Consequently, the substantial variation in biomass among these grouped genotypes constrains high predictive accuracy. Furthermore, we recognize that the genotype-by-environment (G × E) interaction exerts significant pressure on final yield, and our simulations were conducted using climate and soil information at a regional rather than a local scale.

Despite these challenges, CNPP exhibited performance comparable to or exceeding those reported in the literature for studies employing similar calibration procedures (e.g., Battisti et al., 2017; Duarte & Sentelhas, 2020). Moreover, the analysis of on-farm data, which is characterized by an operational rather than a more rigorous experimental approach, suggested the module's suitability for predicting crop yield especially for maize at a regional scale. For soybean, performance was comparatively lower, but still acceptable when compared to studies in the literature that simulates in large

scale (Singh et al., 2024). This regional predictive capacity is of considerable importance for integration with soil organic carbon (SOC) models, indicating CNPP's potential to provide acceptable estimates of carbon input across the evaluated regions.

While parametrizing a crop model often relies on flux tower data, which provides high-resolution measurements of ecosystem exchanges, an alternative parametrization strategy must be employed when such data are unavailable. Unlike the case for maize and soybean, the species parameters for wheat, common bean, and perennial forage had to be calibrated without flux tower data. Instead, parameters governing species-specific processes were adjusted based on established literature values. Despite the loss of accuracy compared to flux tower-assisted calibration, this approach offers significant practical advantages by eliminating the need for resource-intensive, long-term field studies, and has demonstrated effectiveness in previous studies (Soltani et al., 2020; Vergel et al., 2022; Zhao et al., 2019).

Wheat exhibits high phenotypic plasticity, growing in both subtropical and tropical regions, and under both rainfed and irrigated conditions. In Brazil, some wheat cultivation occurs in tropical regions, including the Neotropical savanna in the central part of the country, but production is concentrated in the southernmost regions, where the rainy season is more accentuated. While this region is generally favorable to crop production, the incidence of rain coinciding with physiological maturity can induce precocious sprouting, resulting in yield reductions (Rigatti et al., 2019). This phenomenon, which is not currently represented in the module, adds uncertainty to the calibration processes.

For common bean, limited soil water data imposed a significant constraint on the simulations. Gaps in the information required to calculate available water—such as sites providing only soil texture or geographic coordinates—necessitated the use of external databases to estimate soil properties and derive water availability.

Meanwhile, for perennial forage the main calibration challenge concerned regrowth. The forage crop tends to restore LAI quickly, thus maximizing its light interception. This regenerative process is stimulated by the removal of apical dominance (if present) and the immediate availability of internal resources, spurring new shoots and leaves to emerge and elongate. But regrowth is an inherently difficult process to model and would have been even more challenging if the harvest had been by animal grazing (Bosi et al., 2020). Although this study did not include animal grazing, the module simulates each "virtual grazing event" as a harvest, analogous to other crops, where compartments are reset immediately post-event. The model did not perform as well as some other models in the literature (Pequeno et al., 2014, 2018). The alternative would be a dedicated grazing process, which may be added later at the cost of some complexity in the model.

The module's inherent simplicity—characterized by simplified physiological processes and a limited parameter set—significantly facilitates calibration and allows for rapid recalibration across diverse scales, including regional, site-specific, genotypic, and management-specific levels. However, this parsimony introduces an inevitable trade-off between model simplicity and the detailed representation of complex crop growth dynamics. In effect, a comprehensive calibration is used to mitigate the impact of oversimplification on predictive accuracy. Our results show that relying solely on yield data for calibration can compromise the calibration, resulting in suboptimal performance. We assert that experimental sites providing more detailed information on biomass compartments (i.e., separate data for vegetative and reproductive parts) offer substantially more support for robust calibration. Unfortunately, these detailed biomass components were not available for the wheat and common bean.

Overall, future efforts should focus on integrating mechanisms and technologies to support the assimilation of regional variability, transitioning the CNPP model from a static tool to a more dynamic monitoring system across large agricultural regions. This context necessitates the adoption of advanced practices in crop modeling and calibration, such as (1) using data assimilation based on remote sensing (Jin et al., 2018), (2) employing hybrid approaches with artificial intelligence (e.g., Chang et al., 2023; Shahhosseini et al., 2021), or (3) using CNPP in an ensemble scheme (e.g., Battisti et al., 2017; Duarte & Sentelhas, 2020). These approaches generally result in higher model accuracy relative to the performance observed in on-farm data using datasets 5, 6, 7, and 8, which are mostly based on operational procedures of harvest.

Carbon markets protocols for agriculture ultimately require models to assess the effect of changes in agricultural practices (e.g., improving crop rotations, adopting no-tillage, crop increased productivity, increased days of soil cover, and avoiding use of fire) on SOC stocks. For the long-term impact of projects on SOC stocks, the CNPP module provides essential information on the impacts of practices on the influx of soil carbon to the litter and, consequently, to the soil, while the SOC dynamics is predicted through its coupled framework (PROCS) in long-term simulations, under diverse climate and management scenarios. PROCS uses carbon inflow information to estimate the SOC turnover, involving calculations of the effects of carbon decomposition on carbon losses and carbon stabilization based on soil properties and management and on climatic variables.

From such dynamics the PROCS-CNPP coupled model can estimate long-term accumulation of SOC, which is the ultimate target of carbon markets. Additionally, a rigorous uncertainty quantification must be performed. This involves several input parameters such as soil, weather, management, and calibration parameters, to determine the full range of potential carbon stock change predicted in the simulations.

Quantifying these uncertainties is mandatory for setting reliable confidence intervals for establishing carbon credits. To demonstrate the functionality of the coupled PROCS-CNPP model, we present results from a multiyear simulation of an intensive cultivation system (Figure S3). All the previously cited requirements were meticulously incorporated into this simulation.

5 | CONCLUSIONS

The model demonstrated strong performance consistency across both the irrigated and rainfed sites within dataset 1, thus confirming the effectiveness of the extended implementation. However, achieving equivalent predictive accuracy proved more challenging for crops lacking long-term experimental datasets, as evidenced by the results for wheat, common bean, and perennial forage. Furthermore, a considerably lower performance was observed during calibrations conducted with on-farm data. This reduction in accuracy is a direct consequence of the unrepresented genotypic and management variability inherent in operational environments, which was not incorporated during the general calibration process. Within this on-farm calibration set, maize exhibited superior performance relative to the other crops. We hypothesize that this superior performance is due to maize's greater capacity to respond in terms of yield across diverse regions, and critically, because the genotype clusters were more consistent (i.e., grouping genotypes with similar maturity characteristics). Such effective clustering was not achieved for the other crop types. Given the model's inherent simplicity, we anticipate that end users can effectively calibrate the module, thereby enabling specific parameter tuning for local conditions to achieve higher predictive accuracy. Consequently, within the major Brazilian maize, soybean, wheat, and common bean production regions considered in this study, CNPP is expected to be an useful tool for estimating carbon sequestration when coupled with the PROCS (Barioni et al., 2025) soil biogeochemical model.

AUTHOR CONTRIBUTIONS

Michel Anderson Almeida Colmanetti: Data curation; formal analysis; investigation; methodology; software; validation; writing—original draft; writing—review and editing. **Henrique Oldoni**: Data curation; formal analysis; investigation; methodology; software; validation; writing—original draft; writing—review and editing. **Leandro Annibal**: Software. **Wagner Wolff**: Data curation; writing—original draft; writing—review and editing. **Rodrigo Pereira Abou Rejaili**: Data curation; writing—original draft; writing—review and editing. **Luis Gustavo Barioni**: Conceptualization; funding acquisition; project administration; supervision; writing—original draft; writing—review and editing. **Ivan Rodrigues de Almeida**: Project administration. **Robert P. Ewing**: Investigation; writing—original draft; writing—review and editing. **Santiago Vianna Cuadra**: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

This work was conducted and supported in the framework of Carbon Project at Scale, a collaborative research project between Embrapa and Bayer CropScience.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Michel Anderson Almeida Colmanetti https://orcid.org/0000-0002-4701-4130

Rodrigo Pereira Abou Rejaili https://orcid.org/0009-0005-4586-1909

REFERENCES

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, *22*(6), 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Amiri, E., Irmak, S., & Araji, H. A. (2022). Assessment of CERES-maize model in simulating maize growth, yield and soil water content under rainfed, limited and full irrigation. *Agricultural Water Management*, *259*, 107271. https://doi.org/10.1016/J.AGWAT.2021.107271

Barioni, L. G., Valladão, B. A., Mourão, V. H. M., Karatay, Y. N., Damian, J. M., Moreno, L. M., & Silva, R. O. (2025). Procarbon-soil—Procs: A dynamic soil carbon model for improved model-data compatibility in carbon farming. *SSRN*. https://doi.org/10.2139/ssrn.4812266

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

Bosi, C., Sentelhas, P. C., Huth, N. I., Pezzopane, J. R. M., Andreucci, M. P., & Santos, P. M. (2020). APSIM-tropical pasture: A model for simulating perennial tropical grass growth and its parameterisation for palisade grass (*Brachiaria brizantha*). *Agricultural Systems*, *184*, 102917. https://doi.org/10.1016/j.agsy.2020.102917

Chang, Y., Latham, J., Licht, M., & Wang, L. (2023). A data-driven crop model for maize yield prediction. *Communications Biology*, *6*(1), 439. https://doi.org/10.1038/s42003-023-04833-y

Cunha, G. R., Pires, J. L. F., Maluf, J. R. T., Pasinato, A., Caierão, E., Silva, M. S., Dotto, S. R., Campos, L. A. C., Felício, J. C., Castro, R. L., Marchioro, V., Riede, C. R., Rosa Filho, O., Tonon, V. D., & Svoboda, L. H. (2006). *Regiões de adaptação para trigo no Brasil* (Circular Técnica 20, Embrapa Trigo). Ministério da Agricultura, Pecuária e Abastecimento.

Dadrasi, A., Torabi, B., Rahimi, A., Soltani, A., & Zeinali, E. (2020). Parameterization and evaluation of a simple simulation model

(SSM-iCrop2) for potato (*Solanum tuberosum* L.) growth and yield in Iran. *Potato Research*, *63*(4), 545–563. https://doi.org/10.1007/s11540-020-09456-y

Duarte, Y. C. N., & Sentelhas, P. C. (2020). Intercomparison and performance of maize crop models and their ensemble for yield simulations in Brazil. *International Journal of Plant Production*, *14*(1), 127–139. https://doi.org/10.1007/s42106-019-00073-5

Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., & Wang, J. (2018). A review of data assimilation of remote sensing and crop models. *European Journal of Agronomy*, *92*, 141–152. https://doi.org/10.1016/j.eja.2017.11.002

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J., & Ritchie, J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy*, *18*(3–4), 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N. G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J. P., Silburn, M., Wang, E., Brown, S., Bristow, K. L., Asseng, S., … Smith, C. J. (2003). An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, *18*(3–4), 267–288. https://doi.org/10.1016/S1161-0301(02)00108-9

Kimball, B. A., Thorp, K. R., Boote, K. J., Stockle, C., Suyker, A. E., Evett, S. R., Brauer, D. K., Coyle, G. G., Copeland, K. S., Marek, G. W., Colaizzi, P. D., Acutis, M., Alimagham, S., Archontoulis, S., Babacar, F., Barcza, Z., Basso, B., Bertuzzi, P., Constantin, J., … Zhou, W. (2023). Simulation of evapotranspiration and yield of maize: An inter-comparison among 41 maize models. *Agricultural and Forest Meteorology*, *333*, 109396. https://doi.org/10.1016/J.AGRFORMET.2023.109396

Longo, M., Jones, C. D., Izaurralde, R. C., Cabrera, M. L., Dal Ferro, N., & Morari, F. (2021). Testing the EPIC Richards submodel for simulating soil water dynamics under different bottom boundary conditions. *Vadose Zone Journal*, *20*, e20142. https://doi.org/10.1002/vzj2.20142

Pequeno, D. N. L., Pedreira, C. G. S., & Boote, K. J. (2014). Simulating forage production of Marandu palisade grass (*Brachiaria brizantha*) with the CROPGRO-Perennial Forage model. *Crop and Pasture Science*, *65*(12), 1335–1348. https://doi.org/10.1071/CP14058

Pequeno, D. N. L., Pedreira, C. G. S., Boote, K. J., Alderman, P. D., & Faria, A. F. G. (2018). Species-genotypic parameters of the CROPGRO Perennial Forage model: Implications for comparison of three tropical pasture grasses. *Grass and Forage Science*, *73*(2), 440–455. https://doi.org/10.1111/gfs.12329

R Core Team. (2024). *R: A language and environment for statistical computing (4.3.3)*. R Foundation for Statistical Computing. https://www.r-project.org/

Rigatti, A., Meira, D., Olivoto, T., Meier, C., Nardino, M., Lunkes, A., Klein, L. A., Fassini, F., Moro, É. D., Marchioro, V. S., & Souza, V. Q. (2019). Grain yield and its associations with pre-harvest sprouting in wheat. *Journal of Agricultural Science*, *11*(4), 142. https://doi.org/10.5539/jas.v11n4p142

Ritchie, J. T., Godwin, D. C., & Otter-Nacke, S. (1985). *CERES-wheat: A user-oriented wheat yield model* (AGRISTARS Publication No. YM-U3- 04442-JSC-18892). Michigan State University.

Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. *Scientific Reports*, *11*(1), 1606. https://doi.org/10.1038/s41598-020-80820-1

Silva, E. H. F. M., Kothari, K., Pattey, E., Battisti, R., Boote, K. J., Archontoulis, S. V., Cuadra, S. V., Faye, B., Grant, B., Hoogenboom, G., Jing, Q., Marin, F. R., Nendel, C., Qian, B., Smith, W., Srivastava, A. K., Thorp, K. R., Vieira Junior, N. A., & Salmerón, M. (2025). Inter-comparison of soybean models for the simulation of evapotranspiration in a humid continental climate. *Agricultural and Forest Meteorology*, *365*, 110463. https://doi.org/10.1016/J.AGRFORMET.2025.110463

Singh, R. S., Singh, K. K., & Gohain, G. B. (2024). Simulating crop yield using the DSSAT v4.7-CROPGRO-soyabean model with gridded weather and soil data. *Modeling Earth Systems and Environment*, *10*(1), 845–853. https://doi.org/10.1007/s40808-023-01807-1

Soltani, A., Alimagham, S. M., Nehbandani, A., Torabi, B., Zeinali, E., Dadrasi, A., Zand, E., Ghassemi, S., Pourshirazi, S., Alasti, O., Hosseini, R. S., Zahed, M., Arabameri, R., Mohammadzadeh, Z., Rahban, S., Kamari, H., Fayazi, H., Mohammadi, S., Keramat, S., … & Sinclair, T. R. (2020). SSM-iCrop2: A simple model for diverse crop species over large areas. *Agricultural Systems*, *182*, 102855. https://doi.org/10.1016/j.agsy.2020.102855

Suyker, A. (2024a). *AmeriFlux BASE US-Ne2 mead—Irrigated maize-soybean rotation site (18–5). AmeriFlux AMP, (dataset)*. https://doi.org/10.17190/AMF/1246085

Suyker, A. (2024b). *AmeriFlux BASE US-Ne3 mead—Rainfed maize-soybean rotation s*ite (18–5). AmeriFlux AMP, (dataset). https://doi.org/10.17190/AMF/1246086

USDA. (1986). *Urban hydrology for small watersheds* (Technical Release 55) (2nd ed.). United States Department of Agriculture. https://www.nrc.gov/docs/ML1421/ML14219A437.pdf

Vergel, A. P. R., Rodriguez, A. V. C., Ramirez, O. D., Velilla, P. A. A., & Gallego, A. M. (2022). A crop modelling strategy to improve cacao quality and productivity. *Plants*, *11*, 157. https://doi.org/10.3390/plants11020157

Zhao, C., Liu, B., Xiao, L., Hoogenboom, G., Boote, K. J., Kassie, B. T., Pavan, W., Shelia, V., Kim, K. S., Hernandez-Ochoa, I. M., Wallach, D., Porter, C. H., Stockle, C. O., Zhu, Y., & Asseng, S. (2019). A SIMPLE crop model. *European Journal of Agronomy*, *104*(August 2018), 97–106. https://doi.org/10.1016/j.eja.2019.01.009

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Colmanetti, M. A. A., Oldoni, H., Annibal, L., Wolff, W., Pereira Abou Rejaili, R., Barioni, L. G., de Almeida, I. R., Ewing, R. P., & Cuadra, S. V. (2026). Calibration and evaluation of a carbon net primary productivity module based on the SIMPLE model. *Agronomy Journal*, *118*, e70309. https://doi.org/10.1002/agj2.70309