in silico Plants, 2024, **6**, diae009
https://doi.org/10.1093/insilicoplants/diae009
Advance access publication 21 June 2024
Original Research

in silico Plants
OXFORD

Original Research

SPECIAL ISSUE: Integrative and Multiscale Modelling VSI

Connecting detailed photosynthetic kinetics to crop growth and yield: a coupled modelling framework

Yufeng He[1,2], Yu Wang[2], Douglas Friedel[3], Meagan Lang[3] and Megan L. Matthews[1,2,*]

[1]Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Champaign, IL 61801, USA
[2]Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Champaign, IL 61801, USA
[3]National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Champaign, IL 61801, USA

[*]Corresponding author's e-mail address: mlmatth2@illinois.edu

Handling Editor: Xin-Guang Zhu

Abstract. Photosynthesis and crop growth are inseparable processes that govern plant carbon assimilation and allocation. An accurate model description of these processes can bridge dynamics at the leaf and canopy levels, assisting in identifying potential photosynthetic improvements that can be converted into increased yield. Integrating multiscale biophysical processes and achieving computational effectiveness for seasonal simulations, however, are challenging. Here, we present a fully coupled modelling framework that integrates a metabolic model of C3 photosynthesis (ePhotosynthesis) and a semi-mechanistic crop growth model (BioCro). We replaced the leaf-level Farquhar photosynthesis model in BioCro with the ePhotosynthesis model that mechanistically describes the photosystem electron transport processes and the C3 carbon metabolism including the Calvin–Benson–Bassham cycle and the photorespiratory pathway. The coupled BioCro-ePhotosynthesis model was calibrated to represent a soybean cultivar and developed to be operationally fast for seasonal simulations. As an example of model application, we conducted a global sensitivity analysis of 26 enzymes under an average daytime intercepted radiation of 400 μmol m^{-2} s^{-1}, identifying 2 enzymes, phosphoglycerate kinase (PGK) and phosphoribulokinase (PRK), which had the largest impact on the leaf-level assimilation. Increasing PGK and PRK by 2-fold was predicted to increase the leaf-level assimilation by 8.3 % and the final seed yield by 6.75 % ± 0.5 % over 4 years of observed field climate data. The coupled BioCro-ePhotosynthesis model provides a seamless and efficient integration between the leaf-level metabolism and the field-level yield over a full growing season. The coupled model could be further applied to investigate non-steady-state photosynthetic processes such as non-photochemical quenching.

KEYWORDS: C3 photosynthesis; crop growth model; field climate conditions; metabolic model; multiscale crop model; soybean yield.

1. INTRODUCTION

There has been growing interest in developing models that can accurately predict crop photosynthesis and yield across different scales, from individual leaves to whole fields and from subseconds to a full growing season (Marshall-Colon *et al.* 2017; Peng *et al.* 2020; Matthews and Marshall-Colón 2021). A multiscale model is essential to connecting these scales and understanding the complex interactions among genotype, environment and management. Although mechanistic models have been developed to describe gene regulation, metabolic pathways, organ development and canopy microclimates, they are often used at their respective scales and thus miss interactions between systems (Long *et al.* 2015). Multiscale models have the potential to provide a direct link between genetic modifications and yield (Wang *et al.* 2019), allowing us to use models to identify engineering targets that are predicted to improve yield and not just gene expression or leaf-level photosynthesis.

Recent studies have used multiscale models to examine the connection between leaf-level photosynthesis with crop yield and to evaluate how changing photosynthesis may impact yield under different climate conditions (Wu *et al.* 2019; Moore *et al.* 2021; He and Matthews 2023). However, the leaf photosynthesis in these models all use a steady-state biochemical model, commonly known as the Farquhar–von Caemmerer–Berry (FvCB) model (von Caemmerer 2000). The simplicity and robustness of the FvCB model allow it to be used in a variety of scenarios. However, the FvCB model does not include the individual photosynthetic reactions and, therefore, is not able to simulate how changing the kinetic properties of these reactions impact assimilation rates. Furthermore, the FvCB model

---

assumes steady-state conditions and is not able to simulate the non-steady-state dynamics that are caused by highly heterogeneous canopy properties and microclimate conditions. In fact, photosynthesis in real-world conditions is rarely at steady-state and there are more than 60 reactions involved in the Calvin–Bensen-Bassham (CBB) cycle, electron transport, photorespiratory metabolism and storage and transport of carbohydrates, starch and sucrose (Zhu *et al.* 2007, 2013; Long *et al.* 2015).

A more detailed model of these photosynthetic reactions was developed using enzymatic kinetics and differential equations (Zhu *et al.* 2007, 2013). This dynamic model was successfully used to identify potential enzymes to target to increase leaf-level photosynthetic rates, which was then validated in transgenic tobacco experiments (Rosenthal *et al.* 2011; López-Calcagno *et al.* 2020). Dynamic metabolic models are useful to identify potential metabolic engineering strategies, but they alone cannot predict how modifying the pathways would impact yield under highly varying field conditions (Matthews and Marshall-Colón 2021). To understand how metabolic engineering strategies will impact crop growth and yield, the detailed metabolic model needs to be coupled with a crop growth model. Such an in-silico model framework would allow system-wide assessment of photosynthesis and yield across scales from seconds to a full growing season, and from individual leaves to canopies.

In this paper, we present a coupled modelling framework that integrates a dynamic model of C3 photosynthesis, ePhotosynthesis (Zhu *et al.* 2007, 2013) and a semi-mechanistic crop growth model, BioCro (Lochocki *et al.* 2022; Matthews *et al.* 2022). The coupled BioCro-ePhotosynthesis model bridges the enzymatic activities of a single leaf with the final crop yield on a field scale. The coupled model is parameterised and calibrated to represent a soybean cultivar. We first show the validity of the model by comparing it with the widely used FvCB model at the leaf level and the canopy level with the standalone BioCro that has the FvCB model incorporated. Using an example of increasing specific CBB enzymatic rates, we demonstrate the predicted impact of the resulting increase in the leaf-level assimilation on the canopy-level assimilation and the final soybean yield. Other use cases for the coupled BioCro-ePhotosynthesis model, including other crop types and non-steady-state photosynthesis simulations, are discussed.

2. METHODS AND MODEL DESCRIPTION

2.1 ePhotosynthesis

ePhotosynthesis is a generic dynamic model that describes C3 photosynthesis processes including metabolic pathways of the CBB Cycle, photorespiration, sucrose and starch synthesis metabolic pathway (Zhu *et al.* 2013). It was recently calibrated to fit the observed photosynthetic parameters, the maximum carboxylation rate of Rubisco (V_{cmax}) and the maximum electron transport and RuBP regeneration rate (J_{max}) for soybean (Shameer *et al.* 2022) using two scaling factors that adjust the maximum activities of several key enzymes [**see Supporting Information**—Table S1]. This approach is based on an assumption that under low internal CO_2 concentration (Ci), Rubisco is the only enzyme limiting photosynthetic rates, while under high Ci, every other enzyme in the CBB cycle may play a limiting role. We adopted this method and calibrated the standalone ePhotosynthesis model against a V_{cmax} (=110 μmol m^{-2} s^{-1}) and J_{max} (=195 μmol m^{-2} s^{-1}) that were measured from the A–Ci curves of field grown soybean (Bernacchi *et al.* 2005). We used the Plantecophys R package (Duursma 2015) to derive the values of V_{cmax} and J_{max} from the ePhotosynthesis model simulated A–Ci curves at steady state. The A–Ci curves were generated by simulations with a light intensity of 1500 μmol m^{-2} s^{-1} and Ci levels of 100, 150, 200, 250, 300, 400, 500, 600, 800, and 1200 ppm. The two scaling factors $\alpha_{Rubisco}$ and α_E [**see Supporting Information**—Table S1] were optimised to be 0.87 and 1.03, respectively.

Following Zhu *et al.* (2013, 2007), the total phosphate concentration in the stroma was assumed to be constant of 30 mmol L^{-1} so that the phosphate is high enough to not be a limiting factor. All enzyme kinetic parameters in the ePhotosynthesis model are for 25 °C. The ePhotosynthesis uses a Q10 function to model the temperature-dependent response for each enzyme reaction rate in the CBB cycle from their rates at 25 °C as $V_{max} = V_{max25} \times Q10^{(T-25)/10}$ (Woodrow and Berry 1988; Kannan *et al.* 2019). For most enzymes, the Q10 parameter was assumed to be 2 unless the literature supported another value [**see Supporting Information**—Table S2].

2.2 Soybean-BioCro

Soybean-BioCro is a canopy-scale crop growth model that is part of the BioCro family of crop models (Lochocki *et al.* 2022; Matthews *et al.* 2022). By default, Soybean-BioCro incorporates the FvCB model of photosynthesis and the Ball-Berry stomatal conductance model to estimate the total canopy assimilation from a 10-layer canopy that contains both sunlit and shaded leaves. Each layer has its own microclimate that includes layer-dependent solar radiation, temperature, relative humidity and leaf area index. Soybean-BioCro partitions this assimilated carbon to the different crop organs based on the photothermal development rate. The growth and senescence of the crop biomasses are estimated at an hourly time step and integrated throughout the growing season (Matthews *et al.* 2022). The model calibration and validation for Soybean-BioCro was conducted against 4 years of observed climate data (i.e. 2002, 2004, 2005 and 2006) with soybean biomass measurements (cv. Pioneer 93B15) collected at the SoyFACE facility (40.04°N, 88.23°W) at the University of Illinois Urbana-Champaign (Morgan *et al.* 2005; Matthews *et al.* 2022).

Here, we updated Soybean-BioCro by separating the total pod biomass from the original version (Matthews *et al.* 2022) into its seed and shell components to more easily compare model predictions with observed yield data that often report the seed biomass only. The carbon partitioning coefficients with added seed components were re-optimised [**see Supporting Information**—Table S3], and the same 4-year data (2002, 2004, 2005 and 2006) were used to evaluate the model performance [**see Supporting Information**—Fig. S1].

2.3 The coupled BioCro-ePhotosynthesis

The ePhotosynthesis model is an ordinary differential equation (ODE)-based dynamic system, originally written in Matlab (Zhu *et al.* 2013). We translated the entire model from Matlab to C++

to improve model performance, particularly the model runtime, and to more easily couple the model with BioCro, which is coded in C++. The ePhotosynthesis C++ model uses the CVODE solver from the Sundials software (Gardner *et al.* 2022) to solve the ODE equations. The coupled Soybean-BioCro and ePhotosynthesis models, referred to as the BioCro-ePhotosynthesis model hereafter, were developed by first compiling ePhotosynthesis as a C++ dynamic library that is directly called as a function in the BioCro's C3 leaf photosynthesis module. We replaced the FvCB model for solving the steady-state leaf-level photosynthesis in BioCro with a function that calls the ePhotosynthesis model and returns the steady-state assimilation. As with the FvCB model, the outputs from the ePhotosynthesis model were coupled with the Ball-Berry stomatal conductance model to find an equilibrium solution of Ci and the net assimilation (An) at the leaf level.

Since the BioCro model is typically run on an hourly time step that is much larger than the time (~60 s) required for a metabolite to reach its steady state (Zhu *et al.* 2007), we mainly used the steady-state assimilation from ePhotosynthesis within BioCro. Although the non-steady-state metabolites were not explicitly output in the current version of the coupled model, the ePhotosynthesis model does preserve all of the metabolite concentrations before a steady state is reached. With our coupled framework, one can easily modify the canopy module to access the non-steady-state concentrations, which could then be used with higher-frequency climate drivers than the current hourly time step.

2.4 Enzyme sensitivity experiments at the leaf level

We used the Latin Hypercube Sampling-Partial Rank Correlation Coefficient (LHS-PRCC) to conduct a sensitivity analysis of multiple enzymes at the leaf level using the ePhotosynthesis model. Both the LHS and PRCC functions are directly available in Matlab. LHS is a highly efficient sampling technique when dealing with large sample sizes (Blower and Dowlatabadi 1994). PRCC was then used to quantify the relative importance of each enzyme to changes in the leaf-level assimilation rate. A total of 26 enzymes in the photorespiration pathway and the CBB cycle [**see Supporting Information**—Tables S1 and S4] were investigated in the sensitivity analyses under the following environmental conditions: photosynthetically active radiation (PAR) = 400 μmol m^{-2} s^{-1}, leaf temperature = 24 °C and Ci = 294 ppm. A relatively low PAR level was used to represent the simulated average daytime radiation received by leaves in the multilayer canopy.

2.5 Simulating full growing seasons

As an example of using the coupled BioCro-ePhotosynthesis model, we simulated full growing seasons for 4 years (2002, 2004, 2005 and 2006), the same years used in the original Soybean-BioCro simulations (Matthews *et al.* 2022). We first compared the canopy assimilation and yield between the coupled BioCro-ePhotosynthesis model and the standalone Soybean-BioCro model to confirm that the coupled model was functioning properly and had similar predictions to the standalone Soybean-BioCro model. We then used the BioCro-ePhotosynthesis model to estimate the yield response when two enzymes identified from the enzyme sensitivity analysis, phosphoglycerate kinase (PGK) and phosphoribulokinase (PRK), were increased two-fold (i.e. 200 %). A 2-fold increase was chosen as the max limit as 2 to 3-fold increases of other CBB enzymes have been experimentally achieved (Ku *et al.* 2000; Simkin *et al.* 2015).

3 RESULTS

The ePhotosynthesis C++ model showed a significant improvement in computational performance over its Matlab counterpart. For a single evaluation of the leaf-level assimilation, the ePhotosynthesis C++ and Matlab versions took about 0.075 and 5 s, respectively [**see Supporting Information**—Table S5]. The time required for a full growing season simulation was approximated to be 3.9 and 260 h, respectively [**see Supporting Information**—Table S5]. The performance of the coupled BioCro-ePhotosynthesis (~6 h) was close to the estimation of the C++ version for a growing season simulation, about 40 times faster than that with the Matlab version. The coupled model took longer to run than the approximation due to how the photosynthesis and stomatal conductance calculations are coupled, which can require multiple calls to the ePhotosynthesis model for a given time step.

3.1 Model consistency at steady-state: ePhotosynthesis verses FvCB and BioCro-ePhotosynthesis verses BioCro

Since the leaf-level net assimilation (An) predicted by ePhotosynthesis was calibrated against the same V_{cmax} and J_{max} used by the FvCB model, their A-Ci curves matched well under the high light condition with an incoming PAR (Q) of 1500 μmol m^{-2} s^{-1} (Fig. 1A). At a lower light level of 400 μmol m^{-2} s^{-1}, the ePhotosynthesis model predicted similar An as the FvCB model at high Ci levels, but predicted slightly higher An (+ 0.33 μmol m^{-2} s^{-1}) at Ci less than 600 ppm (Fig. 1B). The light level of 400 μmol m^{-2} s^{-1} was selected for this comparison because the average simulated daytime radiation intercepted by the soybean leaves in Soybean-BioCro's multilayer canopy was approximately this intensity. Overall, the two models showed small differences in An at different light intensity levels with an average of 0.77 μmol m^{-2} s^{-1} higher predicted by the ePhotosynthesis model (Fig. 1C). There were also small differences in the temperature responses of the two models at a 400 μmol m^{-2} s^{-1} light intensity (Fig. 1D). A larger difference between the model predictions occurred when the temperatures became less than 20 °C, with the ePhotosynthesis model predicting an An about 2.7 μmol m^{-2} s^{-1} greater than the FvCB model. For leaf temperatures greater than 20 °C, however, the difference between the two models was less pronounced (Fig. 1D). With an average air temperature at our study site of approximately 22 ± 5 °C over the growing season, the difference at lower temperatures had little impact.

At the canopy level, the coupled BioCro-ePhotosynthesis model predicted slightly higher An than the standalone BioCro. The cumulative An at the end of the growing season was 12.9 t/ha for BioCro-ePhotosynthesis and 12.3 t/ha for BioCro (Fig. 2A). Little difference was observed at the early growing season (Fig. 2B,C) and larger differences occurred at the late growing season as the difference accumulated over time (Fig. 2D). The difference was

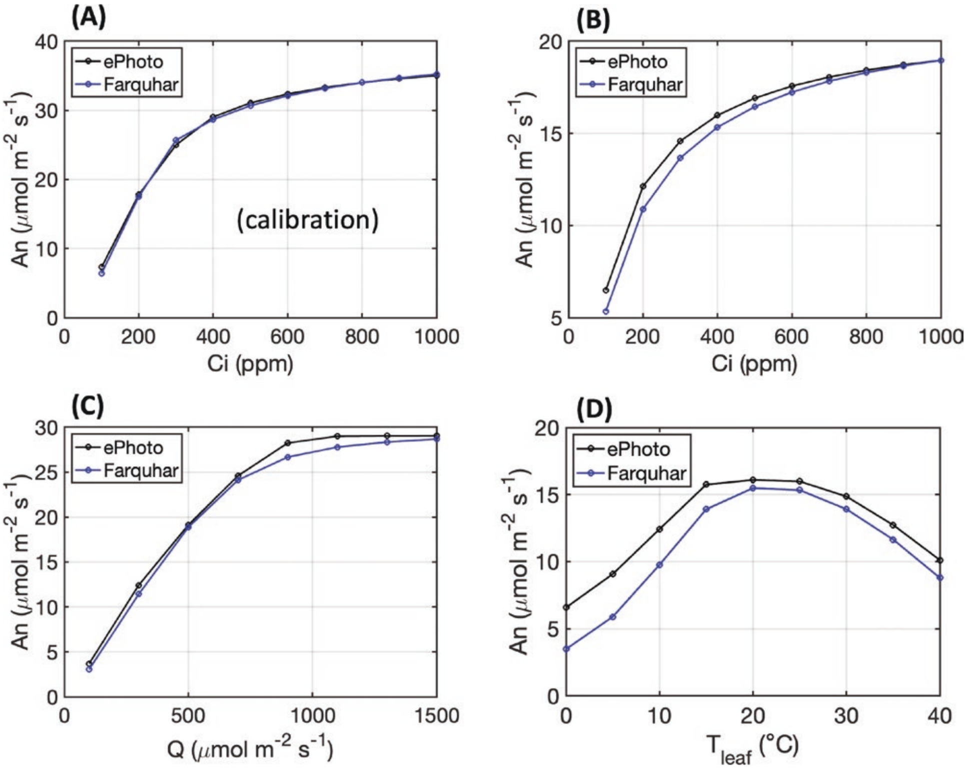

Figure 1. Comparing the leaf-level ePhotosynthesis and FvCB models. The FvCB model has V_{cmax} of 110 and J_{max} of 195 μmol m^{-2} s^{-1}. An represents the leaf-level assimilation. (A) *A*–Ci curves at leaf temperature (Tleaf) of 25 °C and incoming light (*Q*) of 1500 μmol m^{-2} s^{-1}. The ePhotosynthesis model was calibrated by fitting its A-Ci curve to match the same V_{cmax} and J_{max} values used by the FvCB model. (B) A–Ci curves at Tleaf of 25 °C and *Q* of 400 μmol m^{-2} s^{-1}. (C) *A*–*Q* curves with Tleaf of 25 °C and Ci of 400 ppm. (D) *A*–*T* curves with *Q* of 400 μmol m^{-2} s^{-1} and Ci of 400 ppm.

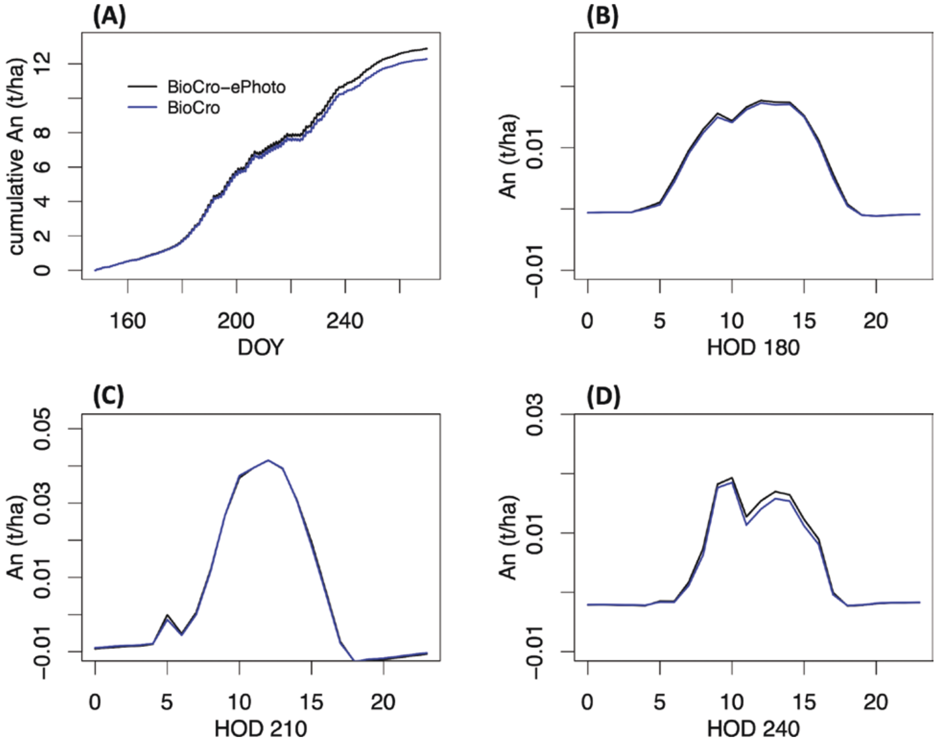

Figure 2. Comparing the canopy-level An between BioCro-ePhotosynthesis and BioCro. (A) cumulative An during the 2006 growing season. (B–D) Hourly diurnal changes of An on the days 180 210, and 240.

also mostly seen during the middle of a day when there was high radiation (Fig. 2D). This difference in the canopy An between the two models originated mostly from the difference in the leaf-level An (Fig. 1). Although a larger leaf area index was predicted by the coupled model, the difference was too small to have a significant impact on the canopy An [**see Supporting Information**—Fig. S2]. Similar to the increase in the canopy An, a slightly increased stomatal conductance was also predicted by the coupled model at the canopy level [**see Supporting Information**—Fig. S3].

3.2 Model application: a 2-fold increase in PGK and PRK concentrations results in a predicted 8.3 % increase in leaf-level assimilation rates

Among the 26 enzyme reactions that were investigated in the sensitivity analysis, reactions V1 (Rubisco), V2 (Phosphoglycerate kinase or PGK), V5 (FBP Aldolase), V9 (Sedoheptulose-bisphosphatase), V10 (Transketolase) and V13 (Phosphoribulokinase or PRK) were found to have the largest influence on An when PAR was set to 400 μmol m^{-2} s^{-1}, Ci was set to 294 ppm and leaf temperature was set to 24 °C (Fig. 3). Increasing the six enzyme concentrations up to 2-fold, all at once and in different combinations, resulted in predicted increases in An up to 9.3 % [**see Supporting Information**—Fig. S4]. Among these enzymes, PGK and PRK were found to be essential targets to achieve an increase in An greater than 8% because they were the only two enzymes that needed to be increased by at least 1.6-fold in all cases [**see Supporting Information**—Fig. S5]. In comparison, there were simulations where any of the other four enzymes could remain at control levels and still result in an 8 % increase in An [[**see Supporting Information**—Fig. S5]. Increasing both PGK and PRK 2-fold in the model led to a predicted 8.3 % increase in An at the leaf level at Ci of ~300 ppm [**see Supporting Information**—Fig. S6A]. Little difference was found between the wildtype and doubled PGK-PRK under the high light condition with PAR of 1500 μmol m^{-2} s^{-1} [**see Supporting Information**—Fig. S6B].

3.3 Model application: a 2-fold increase in PGK and PRK concentrations results in a predicted 7 % increase in soybean yield

Using the coupled BioCro-ePhotosynthesis model, we simulated the impacts of the 2-fold increase in the PGK and PRK

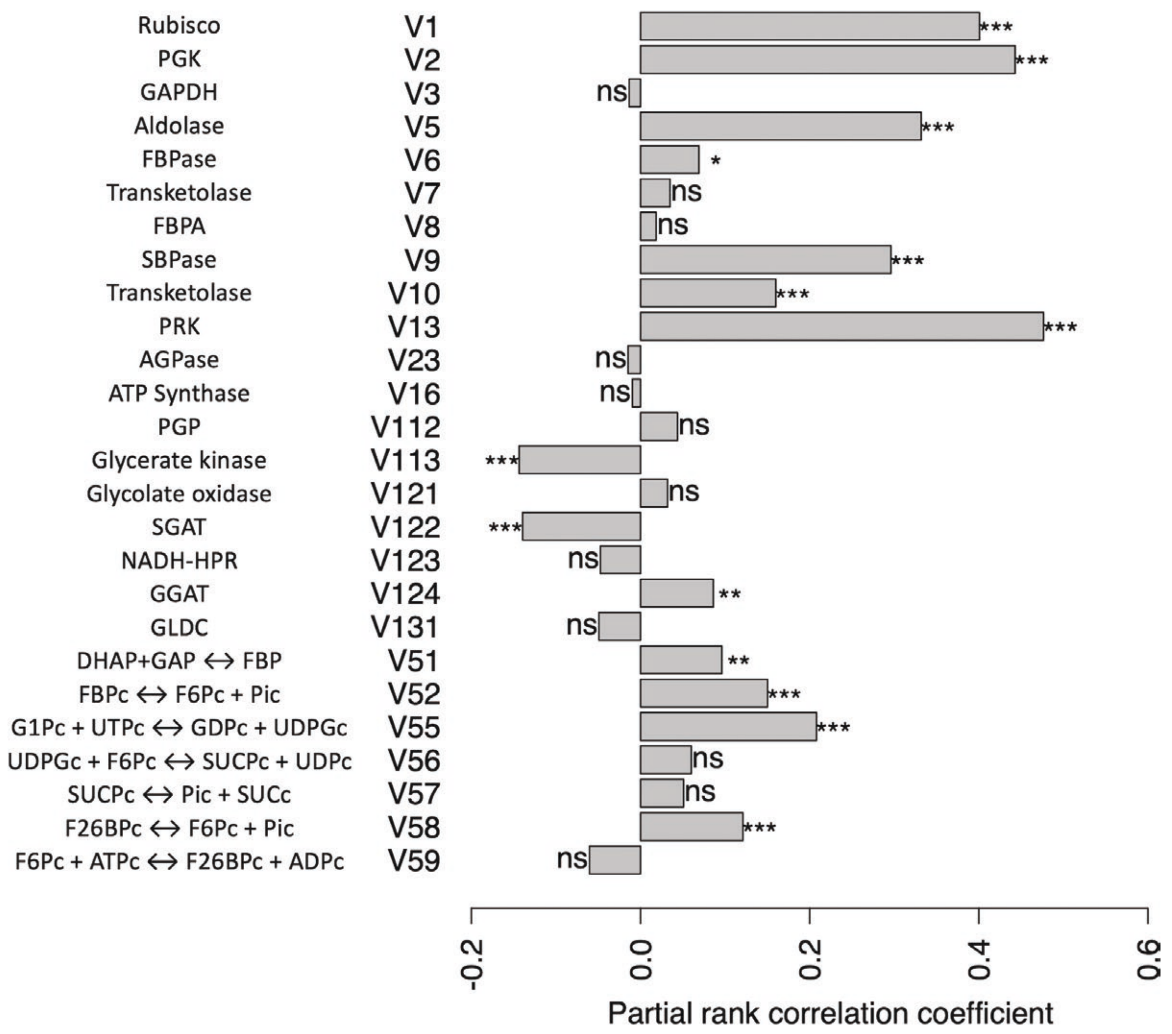

Figure 3. LHS-PRCC of key enzyme reactions in the photorespiration and CBB cycle pathways. The driving environmental conditions were, PAR = 400 μmol m^{-2} s^{-1}, Ci = 294 ppm and leaf temperature = 24 °C. The description of the enzyme reactions' abbreviations can be found in **Supporting Information**—Table S6.

enzyme concentrations on crop growth over four growing seasons, 2002, 2004, 2005 and 2006. The cumulative canopy An at the season end of the 2006 growing season was predicted to be 12.9 t/ha for the wildtype and 13.6 t/ha when PGK and PRK were increased 2-fold (Fig. 4A). The increases in An were observed throughout the season and were the most significant during the middle of the day when the canopy experienced high light (Fig. 4B–D). There was also a notable increase in canopy stomatal conductance observed during midday, coinciding with elevated levels of PGK and PRK enzymes [**see Supporting Information**—Fig. S7].

Larger biomasses were further observed in all the soybean organs, including leaf, stem, seed and shell (Fig. 5). The wildtype predicted seed yield was 5.4 t/ha. When PGK and PRK were increased 2-fold, the predicted yield increased to 5.8 t/ha, a 7 % yield increase. Similar yield increases were also observed in the other years ranging from a 6.2 % increase in 2004 to a 7.3 % increase in 2002 (Fig. 6). On average, increasing the concentration of the PGK and PRK enzymes two-fold resulted in a 6.75% ± 0.5 % predicted increase in seed yield over the four growing seasons.

4 DISCUSSION

4.1 The coupled BioCro-ePhotosynthesis model can efficiently simulate a detailed metabolic pathway at the field-scale over full growing seasons

Crop photosynthesis and growth are complex processes. Simulating their interactions requires a fully coupled model to bridge processes at different scales. The coupled model framework presented here incorporates a detailed kinetic C3 photosynthesis model (ePhotosynthesis) and a crop growth model (BioCro), which allows a seamless and efficient connection between the leaf-level metabolic activities that occur over seconds and the field-level crop growth and yield production that occurs over a full growing season. By translating the ePhotosynthesis model to C++, the coupled model is able to simulate a full growing season within an operationally useful time frame of about 6 h [**see Supporting Information**—Table S5]. Performance could be further sped up in the future by using parallelisation strategies to distribute the multilayer canopy photosynthesis calculations across multiple CPU cores.

We demonstrated one application of the model, where changing the concentration levels of two enzymes led to an increase in the final predicted yield. Previous studies often conducted experiments under high-light conditions. Our sensitivity analysis was under a rather low light condition with limited adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH) production. Therefore, the two enzymes, PGK and PRK (V2 & V13) became predominated since both enzymes are related to ATP consumption. Both PGK and PRK have been found to impact photosynthesis and growth in C3 crops (Paul *et al.* 1995; Rosa-Téllez *et al.* 2018). Although increasing PGK and PRK specifically has yet to be explored, previous studies have shown the potential of overexpressing them, possibly combined with other enzymes in the CBB cycle, to increase photosynthesis (Raines 2022; Yin *et al.* 2022). SBPase and Aldolase (V9 & V5) have received greater attention in previous research (Rosenthal *et al.* 2011; Simkin *et al.* 2015) in the context of another C3 crop, tobacco. The two enzymes were also found to be significant, mostly requiring an increase of at least 1.3-fold [**see Supporting Information**—Fig. S5]. In contrast,

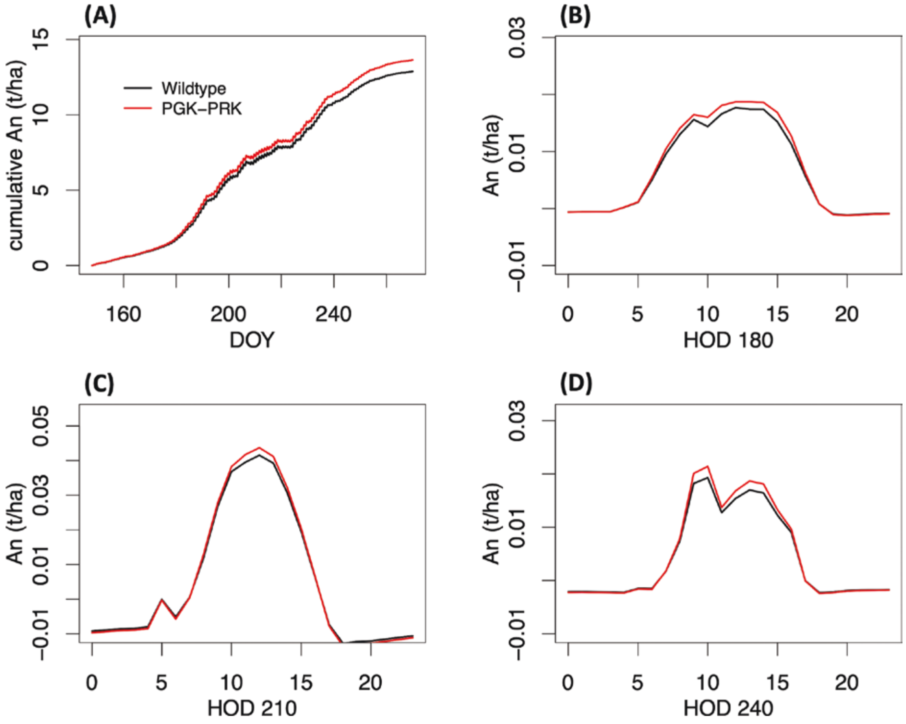

Figure 4 Comparing canopy An for the 2006 growing season between the wildtype and increased PGK and PRK simulations. (A) Cumulative An throughout the growing season. (B–D) Hourly diurnal changes of An on the days 180 210, and 240.

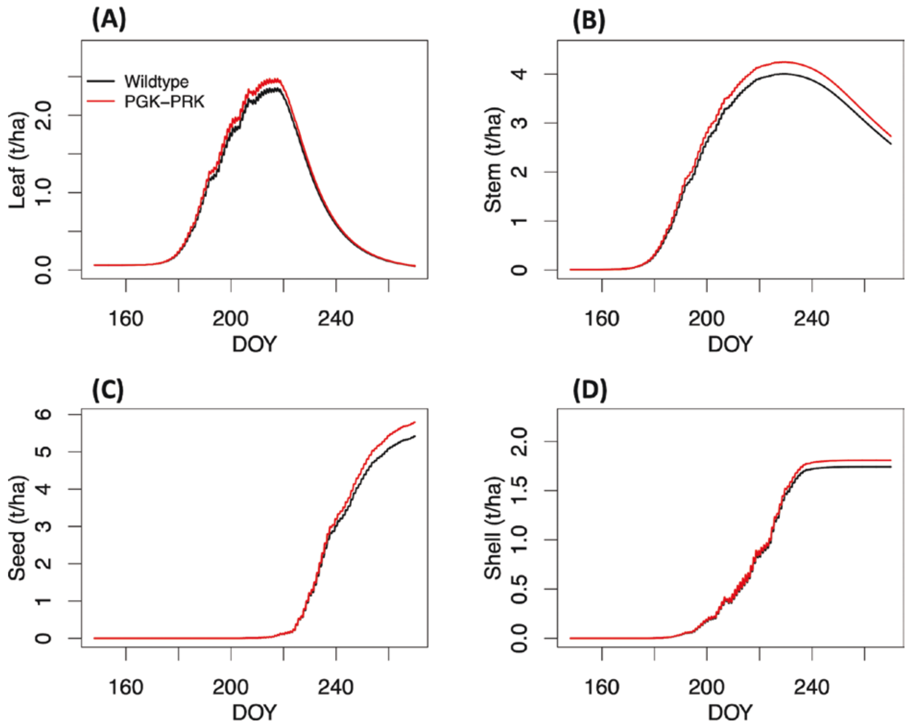

Figure 5. Comparing the wildtype and increased PGK and PRK predicted (A) leaf, (B) stem, (C) seed and (D) shell biomasses for the 2006 growing season.

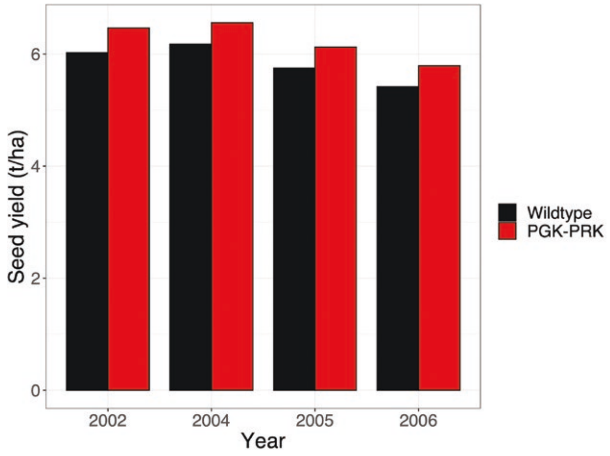

Figure 6. Model predicted wildtype and increased PGK and PRK soybean seed yields for the simulated 2002, 2004, 2005 and 2006 growing seasons.

Rubisco (V1) and Transketolase (V10) exhibited more uniform distributions [**see Supporting Information**—Fig. S5].

4.2 The coupled model can be easily expanded to simulate other C3 and C4 crops

The coupled model can be applied to other crops without much effort needed for additional model development. Both ePhotosynthesis and BioCro were designed to be adjustable for generic crop types. In fact, ePhotosynthesis has been applied to several C3 crops including tobacco (Rosenthal *et al.* 2011), soybean (Shameer *et al.* 2022) and cassava (De Souza *et al.* 2020), and has also been expanded to model C4 crops including sorghum and maize (Wang *et al.* 2021). BioCro has also been used to simulate the growth of several C3 and C4 crops including soybean (Matthews *et al.* 2022), Miscanthus (Miguez *et al.* 2009; He *et al.* 2022) and sugarcane (Jaiswal *et al.* 2017). Any of these crops could be simulated with some parameter changes using this coupled BioCro-ePhotosynthesis framework.

4.3 The importance of evaluating the rate-limiting reactions at canopy level

We demonstrated the general sensitivity of leaf-level An to changes in 26 enzyme concentrations under a low-light environmental condition (Fig. 3). The sensitivity results led us to double the concentrations of PGK and PRK, causing increases in leaf-level An and further in canopy An and yield (Figs. 4–6). There is not a one-to-one relationship between the increase in leaf-level An (8.3 %) to the increase in the cumulative canopy An (5.75 % ± 0.38 %) or the final yield (6.75 % ± 0.5 %), demonstrating the importance of using multiscale models to connect improvements at the leaf-level to the field-scale crop outputs.

Although the leaf-level sensitivity is useful in identifying potential enzymes that may eventually transfer to improvements at the canopy level, it may not be the most optimal choice since the canopy has more heterogeneity in its microclimate and leaf properties. For example, light received by the bottom and top leaves differ (He and Matthews 2023), and considering only an average daytime light for the entire canopy in the sensitivity analysis may be an oversimplification. In addition, the incoming radiation intercepted by the different parts of the canopy changes rapidly within a day, and the reactions that are limiting An may be different under these varying conditions. Further, water and nitrogen availability can also impact photosynthesis and yield at different stages of crop growth (Wu *et al.* 2023). The soybean-BioCro model used in this study assumed abundant water and nitrogen conditions, which may be reasonable for the specific site but could overestimate yield gain in locations with less favourable conditions. Future work with the BioCro-ePhotosynthesis model will include conducting a more comprehensive sensitivity analysis to identify the limiting enzymatic reactions at the canopy level, which may further improve the final yield.

4.4 Use the coupled model to incorporate non-steady-state photosynthesis at the field-scale

Simulating non-steady-state photosynthesis is another key application of the coupled model. Recent advances in non-steady-state photosynthesis focus mainly on the photosynthetic response to fluctuating light (Long *et al.* 2022), which is related to a protective mechanism in plants known as non-photochemical quenching (NPQ). Existing models have been developed to represent the dynamics of chlorophyll fluorescence excitation, electron transport and heat dissipation in the photosystem II supercomplex (Zaks *et al.* 2012; Zhu *et al.* 2013). However, similar to the standalone ePhotosynthesis they are not capable of directly evaluating the impact of NPQ on the canopy CO_2 assimilation. The coupled BioCro-ePhotosynthesis model provides a more comprehensive model framework to integrate NPQ into a full crop photosynthesis and growth model.

While the coupled model is set up to be able to handle non-steady-state dynamics, some challenges remain to using it to study the impacts of non-steady-state photosynthesis at the field-scale. One challenge is the need for high-frequency environmental data, particularly light interception throughout the canopy. The high-frequency inputs would significantly increase the computation time, especially for an entire growing season. Efforts to improve the parallelisation of the canopy photosynthesis calculations to speed up computation time could address this challenge. Another related challenge is the need for a detailed radiation and canopy structural representation over a growing season that captures the high frequency light intensity interception by individual leaves. There have been several efforts in recent years to develop such models (Song *et al.* 2017; Acevedo-Siaca *et al.* 2020; Wang *et al.* 2020; Zhen *et al.* 2022), but more work advancing field light measurements, ray-tracers and high-performance computing resources are needed before these models would be suitable for examining non-steady-state dynamics at the field-scale over a growing season.

4.5 Towards a flexible multidisciplinary model framework

The development of the coupled BioCro-ePhotosynthesis model has involved a range of expertise in biology from cell to whole plant levels, as well as in climate and computer science. The BioCro-ePhotosynthesis model framework can be used to investigate dynamics at the metabolite level and connect to crop growth at the canopy level. However, the users will still need to tailor the model to their own needs by conducting new parameterization, adding functions and extracting necessary model information. The coupled BioCro-ePhotosynthesis model could be further extended to include new plants, other metabolic processes, regulatory pathways or additional ecosystem dynamics. A collective community effort is crucial to advance, develop and apply the model to answer the many open questions around engineering crops for the future.

SUPPORTING INFORMATION

The following additional information is available in the online version of this article –

Figure S1. Model calibration and validation of the updated Soybean-BioCro that separates the pod biomass into the seed and shell biomasses. Four years of data were used: (A) 2002, (B) 2004, (C) 2005 and (D) 2006. Points are the observed data at ambient CO_2 (372 ppm) and lines are the model results. The carbon partitioning coefficients with added seed components were re-optimized using years 2002 and 2005 following the same method used in the original study (Matthews *et al.* 2022). The same four year data were then used to evaluate the model performance. The validation of the original Soybean BioCro version with the pod component can be found in (Matthews *et al.* 2022).

Figure S2. (A) Comparing the leaf area index (LAI) between the BioCro-ePhotosynthesis and the standalone BioCro for the 2006 growing season. (B) Comparing the hourly LAI between the BioCro ePhotosynthesis and the standalone BioCro on DOY 240. (C) Comparing the hourly LAI during the daytime on DOY 240 among three simulations: (1) BioCro-ePhotosynthesis; (2) the standalone BioCro; and (3) a BioCro simulation with the initial state of soybean (e.g. larger LAI) at the last hour of DOY 239 by the BioCro-ePhotosynthesis (dashed red line). This test suggests that the slightly larger LAI only contributed to a small part of the total difference we observed in the canopy An on DOY 240 (see Fig. 2D).

Figure S3. Comparing the canopy-level stomatal conductance between BioCro-ePhotosynthesis and BioCro. (A) cumulative canopy conductance during the 2006 growing season. (B–D) Hourly diurnal changes of canopy conductance on the days 180, 210, and 240.

Figure S4. Frequency distribution of the changes in An (%) due to the changes in six enzyme concentrations associated with reactions V1, V2, V5, V9, V10 and V13. Each enzyme was scaled up by factors of 1, 1.3, 1.6, and 1.9, resulting in a total combination of 4096. Enzyme descriptions can be found in Table S1. Red-circled bars represent the An changes that are larger than 8%, totaling 408 combinations.

Figure S5. Frequency distributions of each scaling factor for the six enzymes that resulted in an increase in An by more than 8% (Fig. S4). Enzyme reaction descriptions can be found in Table S1.

Figure S6. Comparing A-Ci curves between the wildtype and doubled PRK and PRK under (A) a low light of 400 $\mu mol\ m^{-2}\ s^{-1}$ and (B) a high light of 1500 $\mu mol\ m^{-2\ s-1}$. The leaf temperature was set to be 25 °C.

Figure S7. Comparing the canopy-level stomatal conductance between the wildtype (black) and increased PGK and PRK (red) simulations. (A) cumulative canopy conductance during the 2006 growing season. (B–D) Hourly diurnal changes of canopy conductance on the days 180, 210, and 240.

Table S1. Scaling factors for the maximum enzyme reaction rates in the ePhotosynthesis model

Table S2. The Q10 values used in the ePhotosynthesis model to describe the temperature response of each enzyme.

Table S3. Comparing the parameters that were re-optimized in the updated version of BioCro. The updated BioCro separated the Pod into Shell and Leaf.

Table S4. In addition to Table S1, descriptions of more enzymes used in the ePhotosynthesis model for the enzyme sensitivity analysis.

Table S5. Estimated computational time needed for the three models: ePhotosynthesis Matlab, ePhotosynthesis C++, and the fully coupled BioCro-ePhotosynthesis. The canopy evaluation and whole-season evaluation are simple estimations from the single evaluation. The canopy evaluation assumes 3 iterations in the FvCB and Ball-Berry coupling and 10 layers each for sunlit and shaded leaves, thus 60 times longer than the single evaluation. The whole season assumes 130 days. Actual computational time may vary with machine configurations.

Table S6. List of abbreviations for enzymes and their reactions (Tables S1 and S4).

ACKNOWLEDGMENTS

All computational simulations were done through the high-performance computation facility of the BioCluster at the Carl R. Woese Institute for Genomic Biology (IGB) at the University of Illinois Urbana-Champaign (UIUC).

SOURCES OF FUNDING

This work was supported by the research project RealizingIncreased Photosynthetic Efficiency (RIPE), funded from 2017-2023 under grant number OPP1172157 by the Bill & Melinda Gates Foundation, Foundation for Food and Agriculture Research, and the U.K. Government's Department for International Development and by Bill & Melinda Gates Agricultural Innovations grant investment ID 57248. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

CONTRIBUTIONS BY THE AUTHORS

Y.H.: Conceptualisation; formal analysis; methodology; software; validation; visualisation; writing—original draft; writing—review and editing. **Y.W.**: Conceptualization; formal analysis; methodology; software; writing—review and editing. **D.F.**: software; methodology; writing—review and editing. **M.L.**: software; methodology; writing—review and editing. **M.L.M.**: Conceptualisation; formal analysis; funding acquisition; methodology; project administration; software; supervision; writing—original draft; writing—review and editing.

CONFLICTS OF INTEREST

None declared.

LITERATURE CITED

Acevedo-Siaca LG, Coe R, Wang Y, Kromdijk J, Quick WP, Long SP. 2020. Variation in photosynthetic induction between rice accessions and its potential for improving productivity. *New Phytologist* 227:1097–1108. doi:10.1111/nph.16454.

Bernacchi CJ, Morgan PB, Ort DR, Long SP. 2005. The growth of soybean under free air [CO_2] enrichment (FACE) stimulates photosynthesis while decreasing in vivo Rubisco capacity. *Planta* 220:434–446. doi:10.1007/s00425-004-1320-8.

Blower SM, Dowlatabadi H. 1994. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *International Statistical Review / Revue Internationale de Statistique* 62:229. doi:10.2307/1403510.

De Souza AP, Wang Y, Orr DJ, Carmo-Silva E, Long SP. 2020. Photosynthesis across African cassava germplasm is limited by Rubisco and mesophyll conductance at steady state, but by stomatal conductance in fluctuating light. *New Phytologist* 225:2498–2512. doi:10.1111/nph.16142.

Duursma RA. 2015. Plantecophys—an R package for analysing and modelling leaf gas exchange data. *PLoS One* 10:e0143346. doi:10.1371/journal.pone.0143346.

Gardner DJ, Reynolds DR, Woodward CS, Balos CJ. 2022. Enabling new flexibility in the SUNDIALS suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software* 48:1–24. doi:10.1145/3539801.

He Y, Jaiswal D, Liang X, Sun C, Long SP. 2022. Perennial biomass crops on marginal land improve both regional climate and agricultural productivity. *GCB Bioenergy* 14:558–571. doi:10.1111/gcbb.12937.

He Y, Matthews ML. 2023. Seasonal climate conditions impact the effectiveness of improving photosynthesis to increase soybean yield. *Field Crops Research* 296:108907. doi:10.1016/j.fcr.2023.108907.

Jaiswal D, De Souza AP, Larsen S, LeBauer DS, Miguez FE, Sparovek G, Bollero G, Buckeridge MS, Long SP. 2017. Brazilian sugarcane ethanol as an expandable green alternative to crude oil use. *Nature Climate Change* 7:788–792. doi:10.1038/nclimate3410.

Kannan K, Wang Y, Lang M, Challa GS, Long SP, Marshall-Colon A. 2019. Combining gene network, metabolic and leaf-level models shows means to future-proof soybean photosynthesis under rising CO2. *In Silico Plants* 1:diz008. doi:10.1093/insilicoplants/diz008

Ku MSB, Ranade U, Hsu T-P, Cho D, Li X, Jiao D-M, Ehleringer J, Miyao M, Matsuoka M, 2000. Photosynthetic performance of transgenic rice plants overexpressing maize C4 photosynthesis enzymes**Citation: Sheehy JE, Mitchell PL, Hardy B, editors. 2000. Redesigning rice photosynthesis to increase yield. In: Proceedings of the Workshop on The Quest to Reduce Hunger: Redesigning Rice Photosynthesis, 30 Nov.-3 Dec. 1999, Los Bafios, Philippines. Makati City (Philippines): International Rice Research Institute and Amsterdam (The Netherlands): Elsevier Science B.V. 293 p. pp. 193–204. doi:10.1016/S0928-3420(00)80015-4.

Lochocki EB, Rohde S, Jaiswal D, Matthews ML, Miguez F, Long SP, McGrath JM. 2022. BioCro II: a software package for modular crop growth simulations. *In Silico Plants* 4:diac003. doi:10.1093/insilicoplants/diac003

Long SP, Marshall-Colon A, Zhu X-G. 2015. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* 161:56–66. doi:10.1016/j.cell.2015.03.019.

Long SP, Taylor SH, Burgess SJ, Carmo-Silva E, Lawson T, De Souza AP, Leonelli L, Wang Y. 2022. Into the shadows and back into sunlight: photosynthesis in fluctuating light. *Annual Review of Plant Biology* 73:617–648. doi:10.1146/annurev-arplant-070221-024745.

López-Calcagno PE, Brown KL, Simkin AJ, Fisk SJ, Vialet-Chabrand S, Lawson T, Raines CA. 2020. Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants* 6:1054–1063. doi:10.1038/s41477-020-0740-1.

Marshall-Colon A, Long SP, Allen DK, Allen G, Beard DA, Benes B, von Caemmerer S, Christensen AJ, Cox DJ, Hart JC, Hirst PM, Kannan K, Katz DS, Lynch JP, Millar AJ, Panneerselvam B, Price ND, Prusinkiewicz P, Raila D, Shekar RG, Shrivastava S, Shukla D, Srinivasan V, Stitt M, Turk MJ, Voit EO, Wang Y, Yin X, Zhu X-G. 2017. Crops in silico: generating virtual crops using an integrative and multi-scale modeling platform. *Frontiers in Plant Science* 8:786. doi:10.3389/fpls.2017.00786.

Matthews ML, Marshall-Colón A. 2021. Multiscale plant modeling: from genome to phenome and beyond. *Emerging Topics in Life Sciences* 5:231–237. doi:10.1042/ETLS20200276.

Matthews ML, Marshall-Colón A, McGrath JM, Lochocki EB, Long SP. 2022. Soybean-BioCro: a semi-mechanistic model of soybean growth. *In Silico Plants* 4:diab032. doi:10.1093/insilicoplants/diab032

Miguez FE, Zhu X, Humphries S, Bollero GA, Long SP. 2009. A semimechanistic model predicting the growth and production of the bioenergy crop Miscanthus × giganteus: description, parameterization and validation. *GCB Bioenergy* 1:282–296. doi:10.1111/j.1757-1707.2009.01019.x.

Moore CE, Meacham-Hensold K, Lemonnier P, Slattery RA, Benjamin C, Bernacchi CJ, Lawson T, Cavanagh AP. 2021. The effect of increasing temperature on crop photosynthesis: from enzymes to ecosystems. *Journal of Experimental Botany* 72:2822–2844. doi:10.1093/jxb/erab090.

Morgan PB, Bollero GA, Nelson RL, Dohleman FG, Long SP. 2005. Smaller than predicted increase in aboveground net primary production and yield of field-grown soybean under fully open-air [CO2] elevation. *Global Change Biology* 11:1856–1865. doi:10.1111/j.1365-2486.2005.001017.x.

Paul MJ, Knight JS, Habash D, Parry MAJ, Lawlor DW, Barnes SA, Loynes A, Gray JC. 1995. Reduction in phosphoribulokinase activity by antisense RNA in transgenic tobacco: effect on CO$_2$ assimilation and growth in low irradiance. *The Plant Journal* 7:535–542. doi:10.1046/j.1365-313X.1995.7040535.x.

Peng B, Guan K, Tang J, Ainsworth EA, Asseng S, Bernacchi CJ, Cooper M, Delucia EH, Elliott JW, Ewert F, Grant RF, Gustafson DI, Hammer GL, Jin Z, Jones JW, Kimm H, Lawrence DM, Li Y, Lombardozzi DL, Marshall-Colon A, Messina CD, Ort DR, Schnable JC, Vallejos CE, Wu A, Yin X, Zhou W. 2020. Towards a multiscale crop modelling framework for climate change adaptation assessment. *Nature Plants* 6:338–348. doi:10.1038/s41477-020-0625-3.

Raines CA. 2022. Improving plant productivity by re-tuning the regeneration of <scp>RuBP</scp> in the Calvin–Benson–Bassham cycle. *New Phytologist* 236:350–356. doi:10.1111/nph.18394.

Rosa-Téllez S, Anoman AD, Flores-Tornero M, Toujani W, Alseek S, Fernie AR, Nebauer SG, Muñoz-Bertomeu J, Segura J, Ros R. 2018. Phosphoglycerate kinases are co-regulated to adjust metabolism and to optimize growth. *Plant Physiology* 176:1182–1198. doi:10.1104/pp.17.01227.

Rosenthal DM, Locke AM, Khozaei M, Raines CA, Long SP, Ort DR. 2011. Over-expressing the C3 photosynthesis cycle enzyme Sedoheptulose-1-7 Bisphosphatase improves photosynthetic carbon gain and yield under fully open air CO$_2$ fumigation (FACE). *BMC Plant Biology* 11:123. doi:10.1186/1471-2229-11-123.

Shameer S, Wang Y, Bota P, Ratcliffe RG, Long SP, Sweetlove LJ. 2022. A hybrid kinetic and constraint-based model of leaf metabolism allows predictions of metabolic fluxes in different environments. *The Plant Journal* 109:295–313. doi:10.1111/tpj.15551.

Simkin AJ, McAusland L, Headland LR, Lawson T, Raines CA. 2015. Multigene manipulation of photosynthetic carbon assimilation increases CO$_2$ fixation and biomass yield in tobacco. *Journal of Experimental Botany* 66:4075–4090. doi:10.1093/jxb/erv204.

Song Q, Wang Y, Qu M, Ort DR, Zhu X. 2017. The impact of modifying photosystem antenna size on canopy photosynthetic efficiency—development of a new canopy photosynthesis model scaling from metabolism to canopy level processes. *Plant Cell and Environment* 40:2946–2957. doi:10.1111/pce.13041.

von Caemmerer S. 2000. *Biochemical models of leaf photosynthesis*. Csiro Publishing.

Wang E, Brown HE, Rebetzke GJ, Zhao Z, Zheng B, Chapman SC. 2019. Improving process-based crop models to better capture genotype×environment×management interactions. *Journal of Experimental Botany* 70:2389–2401. doi:10.1093/jxb/erz092.

Wang Y, Burgess SJ, de Becker EM, Long SP. 2020. Photosynthesis in the fleeting shadows: an overlooked opportunity for increasing crop productivity? *The Plant Journal: for Cell and Molecular Biology* 101:874–884. doi:10.1111/tpj.14663.

Wang Y, Chan KX, Long SP. 2021. Towards a dynamic photosynthesis model to guide yield improvement in C4 crops. *The Plant Journal: for Cell and Molecular Biology* 107:343–359. doi:10.1111/tpj.15365.

Woodrow IE, Berry JA. 1988. Enzymatic regulation of photosynthetic CO$_2$, fixation in C3 plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 39:533–594. doi:10.1146/annurev.pp.39.060188.002533.

Wu A, Brider J, Busch FA, Chen M, Chenu K, Clarke VC, Collins B, Ermakova M, Evans JR, Farquhar GD, Forster B, Furbank RT, Groszmann M, Hernandez-Prieto MA, Long BM, Mclean G, Potgieter A, Price GD, Sharwood RE, Stower M, van Oosterom E, von Caemmerer S, Whitney SM, Hammer GL. 2023. A cross-scale analysis to understand and quantify the effects of photosynthetic enhancement on crop growth and yield across environments. *Plant Cell and Environment* 46:23–44. doi:10.1111/pce.14453.

Wu A, Hammer GL, Doherty A, von Caemmerer S, Farquhar GD. 2019. Quantifying impacts of enhancing photosynthesis on crop yield. *Nature Plants* 5:380–388. doi:10.1038/s41477-019-0398-8.

Yin X, Tang M, Xia X, Yu J. 2022. BRASSINAZOLE RESISTANT 1 mediates brassinosteroid-induced calvin cycle to promote photosynthesis in tomato. *Frontiers in Plant Science* 12:811948. doi:10.3389/fpls.2021.811948

Zaks J, Amarnath K, Kramer DM, Niyogi KK, Fleming GR. 2012. A kinetic model of rapidly reversible nonphotochemical quenching. *Proceedings of the National Academy of Sciences of the United States of America* 109:15757–15762. doi:10.1073/pnas.1211017109.

Zhen S, van Iersel MW, Bugbee B. 2022. Photosynthesis in sun and shade: the surprising importance of far-red photons. *The New Phytologist* 236:538–546. doi:10.1111/nph.18375.

Zhu X-G, de Sturler E, Long SP. 2007. Optimizing the distribution of resources between enzymes of carbon metabolism can dramatically increase photosynthetic rate: a numerical simulation using an evolutionary algorithm. *Plant Physiology* 145:513–526. doi:10.1104/pp.107.103713.

Zhu X-G, Wang Y, Ort DR, Long SP. 2013. *e*-photosynthesis: a comprehensive dynamic mechanistic model of C3 photosynthesis: from light capture to sucrose synthesis. *Plant Cell and Environment* 36:1711–1727. doi:10.1111/pce.12025.