



OPEN ACCESS

EDITED BY

Paul P. G. Gauthier,
The University of Queensland, Australia

REVIEWED BY

Titta Katariina Kotilainen,
Natural Resources Institute Finland (Luke),
Finland

Inigo Auzmendi,
The University of Queensland, Australia
Isaac Uyehara,
Max Planck Institute of Animal Behaviour,
Germany

*CORRESPONDENCE

André Bénard

✉ benard@msu.edu

RECEIVED 23 November 2022

ACCEPTED 14 April 2023

PUBLISHED 23 May 2023

CITATION

Abedi M, Tan X, Stallknecht EJ, Runkle ES,
Klausner JF, Murillo MS and Bénard A
(2023) Incorporating the effect of
the photon spectrum on biomass
accumulation of lettuce using
a dynamic growth model.
Front. Plant Sci. 14:1106576.
doi: 10.3389/fpls.2023.1106576

COPYRIGHT

© 2023 Abedi, Tan, Stallknecht, Runkle,
Klausner, Murillo and Bénard. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Incorporating the effect of the photon spectrum on biomass accumulation of lettuce using a dynamic growth model

Mahyar Abedi¹, Xu Tan¹, Eric J. Stallknecht², Erik S. Runkle²,
James F. Klausner¹, Michael S. Murillo³ and André Bénard^{1*}

¹Department of Mechanical Engineering, Michigan State University, East Lansing, MI, United States,

²Department of Horticulture, Michigan State University, East Lansing, MI, United States, ³Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI, United States

Cultivation studies in specialty crop optimization utilize models to estimate the fresh and dry mass yield. However, the spectral distribution and photon flux density ($\mu\text{mol m}^{-2} \text{s}^{-1}$) affect plant photosynthetic rate and morphology, which is usually not incorporated in plant growth models. In this study, using data for indoor-grown lettuce (*Lactuca sativa*) cultivated under different light spectra, a mathematical model that incorporates these effects is presented. Different experimental cases are used to obtain a modified quantum use efficiency coefficient that varies with the spectral distribution. Several models for this coefficient are fitted using experimental data. Comparing the accuracy of these models, a simple first- or second-order linear model for light-use efficiency coefficient has about 6 to 8 percent uncertainty, while a fourth-order model has a 2 percent average error in prediction. In addition, normalizing overall spectral distribution leads to a more accurate prediction of the investigated parameter. A novel mathematical model based on normalized spectral irradiance integrated over wavelength for photosynthetically active radiation (PAR) wavebands and the far-red waveband is presented in this study. It accurately predicts lettuce dry mass grown indoors under different light spectra.

KEYWORDS

plant growth, dynamic modeling, spectral distribution, *Lactuca sativa*, indoor crop production, regression-based modeling, controlled environment agriculture

1 Introduction

The photon flux density and spectrum can independently and interactively affect crop photosynthesis, secondary metabolism, and other physiological processes (Ooms et al., 2016). Paz et al. (2019) investigated the impact of DLI (i.e., daily light integral), varying from 1.6 to 9.7 $\text{mol m}^{-2} \text{day}^{-1}$ on the growth of indoor-grown red-leaf lettuce and suggested a minimum DLI of 6.5 $\text{mol m}^{-2} \text{day}^{-1}$. However, biomass of lettuce continues to

increase with DLI until some saturating value, when appearance of physiological disorders begin to appear (e.g., around 17 mol m⁻² day⁻¹) (Both et al., 1994; Kelly et al., 2020). While it is common for crops to have species- and cultivar-specific DLI recommendations for maximized growth rate, the spectral distribution at a constant DLI has additional impacts on biomass accumulation and morphology. For example, decreasing the red to far-red ratio (R:FR) typically increases extension growth (e.g., greater leaf area or elongated stems) that often increases per-plant biomass as a result of increased photon interception (Park and Runkle, 2018a; Park and Runkle, 2018b; Park and Runkle, 2019; He et al., 2021). Similarly, increasing the fraction of blue (B) light a plant receives inhibits extension growth and light interception and can decrease the per-plant biomass of lettuce (Meng et al., 2019; Park and Runkle, 2019; Kong and Nemali, 2021). Increasing the fraction of B and UV light can also increase the biosynthesis of secondary metabolites like anthocyanins which act as photo-protectants and can influence the photosynthetic rate (Meng and Runkle, 2019; He et al., 2021). As an additional consideration to the effect of light intensity and spectral distribution, PAR does not have a constant quantum yield of photosynthesis (mol CO₂ assimilated per mol photon absorbed) on a per-nanometer basis; red light typically has a greater quantum yield than blue or green light (Hogewoning et al., 2012). Meng et al. (2020) analyzed the interaction of blue and green light on hydroponic lettuce growth, and replacing green with red light increased the quantum yield of photosynthesis.

Innovations have been made with respect to spectral-shifting materials for agricultural use that attempt to leverage our understanding of how light intensity and spectrum influence crops. For example, Shen et al. (2021) developed a spectral-shifting film that primarily absorbs blue and green light and fluoresces red and far-red light to theoretically increase lettuce biomass accumulation through increased quantum efficiency and light interception. Hebert et al. (2022) constructed luminescent quantum dot films that decrease overall DLI by 14%, but the modified spectrum enhances the tomato biomass yield and vegetative growth by 6% and 10%, respectively. Despite the wealth of knowledge on how light intensity and distribution affect crop growth, models predicting crop growth have not developed at a similar rate.

A plant growth model is a valuable tool to predict yield and provide an approximation for the impact of factors (e.g., water use or CO₂ concentration). In addition, plant growth modeling allows researchers to perform virtual studies to test a hypothesis without investing the required time to perform costly experiments. Van Henten and Van Straten (1994) developed a dynamic model to predict lettuce dry mass as a state variable in time using environmental inputs such as CO₂ concentration, spectral irradiance for photosynthetically active radiation, and ambient temperature. Jones et al. (1991) proposed a model for tomato growth that responds to constantly varying environmental parameters, and the plant state was presented through seven variables that included dry mass for different components, leaf number, and leaf area. These are two of several computational models that consider environmental parameters to increase crop yield. While these models consider the impacts of spectral

distribution through the overall spectral irradiance (overall energy of the incoming spectrum), the impacts of spectral distribution on a photometric basis are often disregarded. There are few models in the literature that incorporate the impact of the photon spectrum of incoming light on plant growth. Dieleman et al. (2019) aimed to investigate the impact of light quality on tomato physiological and morphological responses. Young tomato plants were cultivated under 7 different light treatments, and various parameters were measured, including leaf light reflection and transmission, accumulated biomass, photosynthesis rate, and concentration of light-capturing pigments. Based on these measurements and the 3D model developed in GroIMP, and when extrapolated to a mature (fruit-bearing crop), it was suggested that dynamic light spectra might stimulate growth and production for an indoor crop production system.

The aim of this study is to modify an existing calibrated dynamic growth model of lettuce to accommodate the impact of spectral distribution. Several regression scenarios are investigated to find a modified model that estimates the impact of spectral distribution and intensity on plant growth. A new modified light-use efficiency coefficient that quantifies the impact of spectral distribution is also presented below.

2 Plant growth computational modeling

2.1 Plant growth model for lettuce

The dynamic growth model of lettuce proposed by Van Henten (1994) is modified in this study to numerically investigate the impact of spectral distribution and intensity on lettuce dry mass and yield. Using dry mass as the primary output for the model, this variable is further subdivided into structural dry mass and nonstructural dry mass, which accounts for starch, glucose, and other similar elements. The model assumes that the two categories of dry mass fully define the state of the plant and describes lettuce growth by calculating these sub-variables using the following ordinary differential equations (ODE),

$$\frac{dX_{nsdm}}{dt} = c_{\alpha} f_{phot} - r_{gr} X_{sdm} - f_{resp} - \frac{1 - c_{\beta}}{c_{\beta}} r_{gr} X_{sdm}, \quad (1)$$

$$\frac{dX_{sdm}}{dt} = r_{gr} X_{sdm}. \quad (2)$$

Equations (1) and (2) represent the transient behavior in the structural and non-structural dry mass per unit of area (g m⁻²) in response to photosynthesis. In the above equations, $f_{phot} = f_{phot}(C_{CO_2}, I, T, X_{sdm})$ is the gross canopy photosynthesis (g m⁻² s⁻¹), $f_{resp} = f_{resp}(T, X_{nsdm}, X_{sdm})$ is the maintenance respiration (g m⁻² s⁻¹), and $r_{gr} = r_{gr}(T, X_{nsdm}, X_{sdm})$ is the growth rate of structural material (g m⁻² s⁻¹), while c_{α} and c_{β} describe the conversion rate of CO₂ to sugar (CH₂O) and yield factor which is a measure of non-structural dry mass losses due to respiration and photosynthetic activities, respectively. The value for c_{α} is the

molecular weight ratio of CO₂ to CH₂O and is set to 0.68. According to Sweeney (1981), c_β for lettuce is approximately 0.8. The growth rate (r_{gr}) refers to the rate at which non-structural materials are transformed into structural materials, i.e,

$$r_{gr} = c_{gr,max} \frac{X_{nsdm}}{c_\gamma X_{sdm} + X_{nsdm}} c_{Q_{10},gr}^{(T-20)/10}, \tag{3}$$

where T is the canopy temperature (°C), $c_{gr,max}$ is the saturation growth rate at 20°C, c_γ is the growth rate coefficient, and $c_{Q_{10},gr}$ is the measure of growth rate sensitivity to the canopy temperature. Van Holsteijn (1981) approximated the saturation growth rate coefficient to $5 \times 10^{-6} \text{ s}^{-1}$; Sweeney (1981) estimated the growth rate coefficient for lettuce to 1.0. The growth rate sensitivity constant is set to 1.6, which means that for every 10°C increase in the canopy temperature, the growth rate increases by a factor of 1.6. The maintenance respiration rate is predicted through,

$$f_{resp} = (c_{resp,sh}(1 - c_\tau)X_{sdm} + c_{resp,rt}c_\tau X_{sdm})c_{Q_{10},resp}^{(T-25)/10}. \tag{4}$$

In Equation (4), $c_{resp,sh}$ and $c_{resp,rt}$ represent shoot and root maintenance respiration coefficients at 25°C and indicate the amount of glucose consumption per structural dry material. Van Henten (1994) estimated shoot and root respiration coefficients as $3.47 \times 10^{-7} \text{ s}^{-1}$ and $1.16 \times 10^{-7} \text{ s}^{-1}$, respectively. $c_{Q_{10},resp}$ is the sensitivity of maintenance respiration to canopy temperature and Van Henten (1994) assigned a value of 2.0 for this coefficient. c_τ is the ratio of root dry mass to the overall dry mass of the plant, which can depend on the type of cultivation. Lorenz and Wiebe (1980) reported an average value of 0.15 for lettuce cultivated in soil, while Sakamoto and Suzuki (2015) measured an average value of 0.14 for hydroponic lettuce cultivation. Goudriaan and Monteith (1990) formulated an empirical correlation to estimate gross canopy photosynthesis,

$$f_{phot} = (1 - \exp(-c_K c_{lar}(1 - c_\tau)X_{sdm}))f_{phot,max}, \tag{5}$$

that c_K is the extinction coefficient, and for lettuce with planophile characteristics, is set to 0.9; c_{lar} is the structural leaf area ratio and Lorenz and Wiebe (1980) approximated it to $75 \times 10^{-3} \text{ m}^2 \text{ g}^{-1}$; and $f_{phot,max}$ is the gross CO₂ assimilation rate for a canopy with 1 square meter of effective surface area. Acock et al. (1978) presented an equation to calculate $f_{phot,max}$ considering the effect of CO₂ concentration and spectral irradiance integrated over wavebands within PAR as well as canopy temperature and photorespiration,

$$f_{phot,max} = \frac{\varepsilon I g_{CO_2} c_\omega (C_{CO_2} - \Gamma)}{\varepsilon I + g_{CO_2} c_\omega (C_{CO_2} - \Gamma)}. \tag{6}$$

In Equation (6), ε is the light-use efficiency, I is the spectral irradiance integrated over the wavebands within PAR that regulates plant growth, g_{CO_2} is the conductance of canopy for the diffusion of CO₂, c_ω is the density of CO₂ that has an approximate value of $1.83 \times 10^{-3} \text{ g m}^{-3}$ (considering greenhouse temperature around 20°C), C_{CO_2} is the concentration of CO₂ in the greenhouse, and Γ is the CO₂ compensation point, accounting for the impact of the temperature on photosynthesis rate (Van Henten, 1994). CO₂ compensation is determined based on canopy temperature using

the following correlation,

$$\Gamma = c_\Gamma c_{Q_{10},\Gamma}^{(T-20)/10}, \tag{7}$$

whereas c_Γ is the CO₂ compensation point at 20°C which is 40 mL L⁻¹, and $c_{Q_{10},\Gamma}$ is the sensitivity of CO₂ compensation with canopy temperature, which Goudriaan et al. (1985) approximated it as 2.0. Light-use efficiency is computed considering light level impact on CO₂ compensation and photorespiration (Goudriaan et al., 1985),

$$\varepsilon = c_\varepsilon \frac{C_{CO_2} - \Gamma}{C_{CO_2} + 2\Gamma}. \tag{8}$$

In Equation (8), c_ε is the quantum use efficiency which is the energy required for a reduction of one mole CO₂, and Goudriaan et al. (1985) approximated its value to be about $17.0 \times 10^{-6} \text{ g J}^{-1}$. In this study, there is an assumption that this parameter is affected by the photon spectral distribution; therefore, its value varies depending on the lighting conditions utilized for lettuce growth. Goudriaan et al. (1985) developed a mathematical correlation for the canopy conductance for CO₂ diffusion, which is derived considering the boundary layer, stomatal, and carboxylation conductance,

$$\frac{1}{g_{CO_2}} = \frac{1}{g_{bnd}} + \frac{1}{g_{stm}} + \frac{1}{g_{car}}, \tag{9}$$

where g_{bnd} , g_{stm} , and g_{car} represent the boundary layer, stomatal, and carboxylation conductance, respectively. Stanghellini (1987) estimated the boundary layer conductance to be 0.007 m s^{-1} at a 5°C temperature gradient, 0.1 m s^{-1} wind speed, and leaf with a characteristic length of 0.075 m. For a plant that grows in an environment without stress, Stanghellini (1987) approximated stomatal conductance to be 0.005 m s^{-1} . Carboxylation conductance is a function of canopy temperature and its value (from 5 to 40°C) is determined using the following empirical correlation,

$$g_{car} = -1.32 \times 10^{-5} T^2 + 5.94 \times 10^{-4} T - 2.64 \times 10^{-3}. \tag{10}$$

Table 1 provides a summary for the definition of different coefficients and their numerical values within the plant growth model.

2.2 Plant growth ODE solver

A MATLAB code was developed to find a solution for the ODE Equations (1), and (2). The code utilized experimental temperature, spectral irradiance integrated over wavebands from 400 to 750 nm, and CO₂ concentration as inputs to compute the two sub-variable dry masses as outputs. Input and output data were extracted from experiments that investigated the impact of the photon spectrum on production of lettuce ‘Rouxai’ growth by Meng and Runkle (2019); Meng et al. (2019), and Meng et al. (2020). Meng and Runkle (2019) carried out three replications with a PPFD of 100 and 180 $\mu\text{mol m}^{-2} \text{ s}^{-1}$ during 0-2 and 2-3 days, respectively. After that, the seedlings were grown under various LED treatments with a 24-hour photoperiod, and the temperature was set to 23°C. Fresh and

TABLE 1 Summary of coefficients needed in Equations (1)-(9) for lettuce cultivation modeling.

| Parameter | Definition | Value | Reference |
|---------------------|--|--|-------------------------------|
| c_{α} | Conversion rate of CO ₂ to CH ₂ O | 0.68 | Van Henten (1994) |
| c_{β} | Yield factor | 0.8 | Sweeney (1981) |
| $c_{gr,max}$ | Saturation growth rate at 20 °C | $5 \times 10^{-6} \text{ s}^{-1}$ | Van Holsteijn (1981) |
| c_{γ} | Growth rate coefficient | 1.0 | Sweeney (1981) |
| $c_{Q_{10},gr}$ | Growth rate sensitivity to the canopy temperature | 1.6 | Sweeney (1981) |
| $c_{resp,sh}$ | Shoot maintenance respiration coefficient at 25 °C | $3.47 \times 10^{-7} \text{ s}^{-1}$ | Van Henten (1994) |
| $c_{resp,rt}$ | Root maintenance respiration coefficient at 25 °C | $1.16 \times 10^{-7} \text{ s}^{-1}$ | Van Henten (1994) |
| $c_{Q_{10},resp}$ | Sensitivity of maintenance respiration to the canopy temperature | 2.0 | Van Henten (1994) |
| c_{τ} | Ratio of root dry mass to total plant dry mass (soil) | 0.15 | Lorenz and Wiebe (1980) |
| | Ratio of root dry mass to total plant dry mass (hydroponic) | 0.14 | Sakamoto and Suzuki (2015) |
| c_K | Extinction coefficient | 0.9 | Goudriaan and Monteith (1990) |
| c_{lar} | Structural leaf area ratio | $75 \times 10^{-3} \text{ m}^2 \text{ g}^{-1}$ | Lorenz and Wiebe (1980) |
| c_{ω} | Density of CO ₂ | $1.83 \times 10^{-3} \text{ g m}^{-3}$ | Van Henten (1994) |
| c_{Γ} | CO ₂ compensation point at 20 °C | 40 mL L^{-1} | Goudriaan et al. (1985) |
| $c_{Q_{10},\Gamma}$ | Sensitivity of CO ₂ compensation with canopy temperature | 2.0 | Goudriaan et al. (1985) |
| c_e | Quantum use efficiency as energy required for a reduction of one molecule of CO ₂ | $17.0 \times 10^{-6} \text{ g J}^{-1}$ | Goudriaan et al. (1985) |
| g_{bnd} | Boundary layer conductance | 0.007 m s^{-1} | Stanghellini (1987) |
| g_{stm} | Stomatal conductance | 0.005 m s^{-1} | Stanghellini (1987) |

dry mass data were obtained using destructive tests for plants harvested on day 10. Similarly, Meng et al. (2019) performed experiments three times at a total photon flux density of $180 \mu\text{mol l m}^{-2} \text{ s}^{-1}$. The seedlings were transplanted into a hydroponic system with a 20-hour photoperiod, an average air temperature of 21.1 °C, average CO₂ concentration of 402 mL L^{-1} , and relative humidity ranging from 41 % to 70%. Meng et al. (2020) conducted two replications at a temperature of 20 °C, a total photon flux density of $50 \mu\text{mol m}^{-2} \text{ s}^{-1}$, and 24-hour photoperiod. The next day, the temperature, the photoperiod, and total photon flux density were set to 22 °C, 20 hours, and $180 \mu\text{mol m}^{-2} \text{ s}^{-1}$, respectively. On the fourth day, the seedlings were exposed to nine different light-quality treatments under the same controlled conditions. For the first replication, the average temperature, relative humidity, and CO₂ concentration was 22.4 °C, 410 mL L^{-1} , and 34 %, and in the second replication, these parameters were 22.4 °C, 410 mL L^{-1} and 35%, respectively. Table 2 represents the experimental data for red-leaf lettuce ‘Rouxai’ under different light treatment conditions, while Figure 1 illustrate the variation of incoming light spectra.

Using integrated spectral irradiance, CO₂ concentration, and temperature, the Van Henten (1994) growth model is used with the parameters in Table 1 to predict dry mass for different light treatment experiments. Figure 2 represents the necessity of considering the impact of spectral distribution on lettuce growth by showing the difference between the experimental data and the plant growth

model. For example, the growth model predicted only 6 of the 20 lighting treatments to be within 25 % of the actual dry mass values.

Numerical error is the measure of a difference between the experimental dry mass of lettuce and the model prediction using the suggested value for c_e (Van Henten, 1994), which was $17.0 \times 10^{-6} \text{ g J}^{-1}$. With the assumption of unvarying c_e for different experiments, the dynamic growth model does not accurately predict for various light treatment experiments other than typical greenhouse light conditions.

2.3 Validation of ODEs solver through altering c_e for different experiments

As mentioned earlier, using a constant c_e led to a considerable error in lettuce dry mass prediction; however, as we will show, a varying c_e depending on spectral distribution allows prediction in good accordance with experimental data. Knowing the dry mass of lettuce for different experiments, the solver tries to find a value for c_e that allows a prediction of a state variable in good accordance with experimental data. These values will be used in the next section to develop a model that predicts the impact of the photon spectrum on c_e , and eventually on plant growth. Figure 3 compares the lettuce dry mass predicted by the model with the experimental dry mass for the investigated light treatments. It is inferred from Figure 3 that the

TABLE 2 Experimental data for different spectral treatments obtained from Meng and Runkle (2019) and Meng et al. (2019, 2020).

| Treatment Number | Literature | Treatment Type | Dry Mass (g) | PFD ($\mu\text{mol m}^{-2}\text{s}^{-1}$) | $I_{\text{PAR+FR}}$ (Wm^{-2}) |
|------------------|------------------------|----------------|--------------|---|--|
| 1 | Meng et al. (2020) | R180 | 2.931 | 180.2 | 32.8 |
| 2 | Meng et al. (2020) | G60R120 | 2.745 | 184.3 | 36.3 |
| 3 | Meng et al. (2020) | B20R160 | 2.258 | 179.6 | 34.4 |
| 4 | Meng et al. (2020) | B20G60R100 | 2.449 | 180.2 | 37.2 |
| 5 | Meng et al. (2020) | B60R120(1) | 2.051 | 183.0 | 38.5 |
| 6 | Meng et al. (2020) | B60G60R60 | 1.527 | 178.7 | 40.3 |
| 7 | Meng et al. (2020) | WW180(1) | 2.470 | 184.8 | 36.5 |
| 8 | Meng and Runkle (2019) | B30R150 | 0.046 | 183.7 | 36.2 |
| 9 | Meng and Runkle (2019) | B30R150FR30 | 0.052 | 216.5 | 41.7 |
| 10 | Meng and Runkle (2019) | R180FR30 | 0.045 | 214.2 | 38.4 |
| 11 | Meng et al. (2019) | B60R120(2) | 1.014 | 178.1 | 37.3> |
| 12 | Meng et al. (2019) | B40G20R120 | 1.187 | 182.5 | 37.3 |
| 13 | Meng et al. (2019) | B20G40R120 | 1.348 | 181.7 | 36.4 |
| 14 | Meng et al. (2019) | G60R120 | 1.587 | 184.3 | 36.3 |
| 15 | Meng et al. (2019) | B40R120FR20 | 1.232 | 180.8 | 36.1 |
| 16 | Meng et al. (2019) | B20R120FR40 | 1.438 | 183.6 | 34.4 |
| 17 | Meng et al. (2019) | R120FR60 | 1.622 | 174.2 | 30.1 |
| 18 | Meng et al. (2019) | B20G20R120FR20 | 1.417 | 182.4 | 35.4 |
| 19 | Meng et al. (2019) | WW180(2) | 1.394 | 188.1 | 37.0 |
| 20 | Meng et al. (2019) | EQW180 | 1.087 | 181.7 | 38.4 |

B, G, R, FR, WW, and EQW refers to blue (400–500nm), green (500–600nm), red (600–700nm), far-red (700–750nm), warm-white, and equalized white light-emitting diodes, respectively according to Meng et al. (2020). PFD, and $I_{\text{PAR+FR}}$ represent photon flux density and spectral irradiance integrated over PAR and far-red wavebands.

solver is capable of finding a quantum use efficiency for each experiment that leads to an accurate prediction of the lettuce state variable (dry mass) on the day of harvest.

3 Implementation of regression methodology to account for the impact of spectral distribution and intensity on lettuce growth

This section describes the development of the linear regression model. Two distinctive datasets are considered, which represent the experiments carried out under different LED spectrums and novel data for natural light. The general form of the model is presented in the next subsection, which is followed by a discussion on the input features of the model. Exploratory data analysis is performed on the input dataset in subsection 3.3. The generic form of the empirical models is investigated in subsection 3.4, and the performance of different models is evaluated using various metrics such as R^2 , mean absolute percentage error (MAPE), Akaike information criterion (AIC), and Bayesian information criterion (BIC), in the next subsection. A regression model is built from the LED lighting data using a train

and test split of 85% and 15%, respectively. The effects of combining the suggested empirical model with the dynamic growth model for lettuce are studied in subsection 3.6. In the last subsection, the precision of the proposed combined dynamic growth model is evaluated using data from another study conducted under completely different experimental conditions under natural lighting.

3.1 Light-use efficiency prediction based on incoming spectrum

The aim of this study is to develop a model that predicts the quantum use efficiency (c_e) as a function of spectral photon flux density or integrated spectral irradiance within the PAR+FR waveband (i.e., 400–750 nm). A linear regression approach proves to be an invaluable tool for generating a basic model for obtaining weights for different features and establishing a simple mathematical model in the form of $c_e = \sum_{i=1}^4 (\omega_i F_i)$ where ω_i and F_i correspond to the weight (coefficient) and the value for the i^{th} input feature, respectively. Since the incoming spectrum is a continuous function (Figure 1), an idea was devised to generate discrete features based on the continuous distribution of the spectrum for the linear

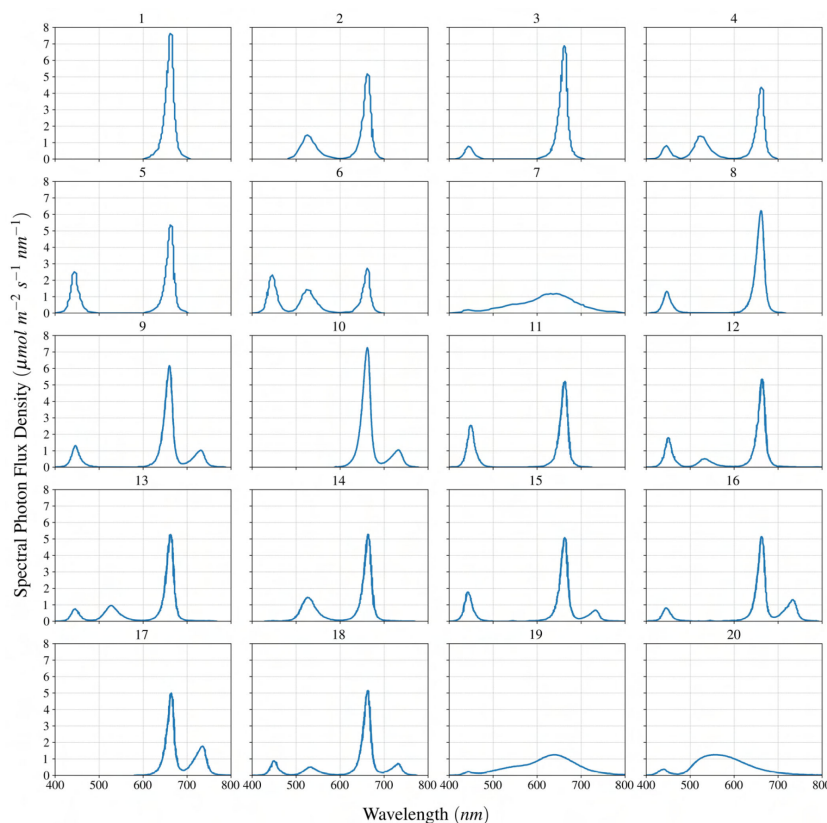


FIGURE 1 Spectral distribution for different case studies of lighting treatment for lettuce reported in the literature to study the effect of spectral distribution on lettuce growth. The label at the top of each graph represents a light treatment experiment according to Table 2.

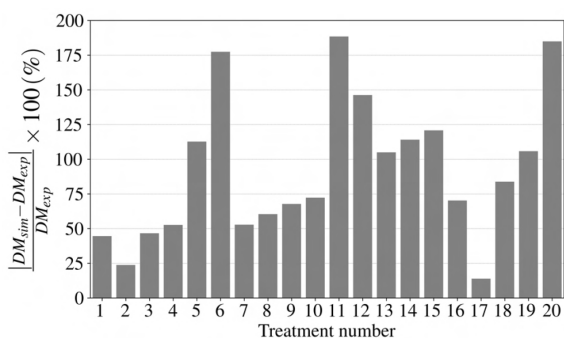


FIGURE 2 Numerical error for the lettuce growth model using the c_e value by Van Henten (1994). DM_{sim} , and DM_{exp} are lettuce dry mass for numerical simulation and experimental study in g, respectively. The considerable differences, some with more than 75 % error, indicate the necessity of considering the impact of spectral irradiance and flux density on the c_e value for an accurate prediction of lettuce growth yield.

regression model, as shown in Figure 4. This idea involves dividing the incoming spectrum into four segments, which in the context of this study, corresponding to four distinctive wavebands in PAR+FR: blue (400-500 nm), green (500-600 nm), red (600-700 nm), and far-red (700-750 nm). A discrete value is assigned for each segment

based on the integral of spectral distribution. Subsection 3.2 provides a detailed description of how these discrete values are obtained for each spectrum.

In addition, it is possible to investigate the interaction between different light wavebands on quantum use efficiency through the regression model in the form of $\bar{c}_e = \sum_{i=1}^4 (\omega_i F_i) + \sum_j (\omega'_j F'_j)$ where F'_j can be defined as the multiplication of two fraction ratios, e.g., the blue and green wavebands. Figure 4 provides an overview of the development of empirical correlation for light-use efficiency based on the continuous spectrum distribution.

3.2 Definition of the fraction ratio

As previously stated, the dynamic growth model's efficiency can be improved by developing a model for c_e that takes into account the impact of spectral distribution. This can be accomplished by creating a function for c_e that is dependent on the photon flux density or integrated spectral irradiance ratio for 100-nm wavebands in PAR and 50-nm FR waveband. Calculation of these ratios based on spectral photon flux distribution is more convenient since Meng et al. (2020) assigned a label based on photon flux density treatments with different wavebands. Considering lighting treatment 18, or "B20G20R120FR20" as an example, the photon

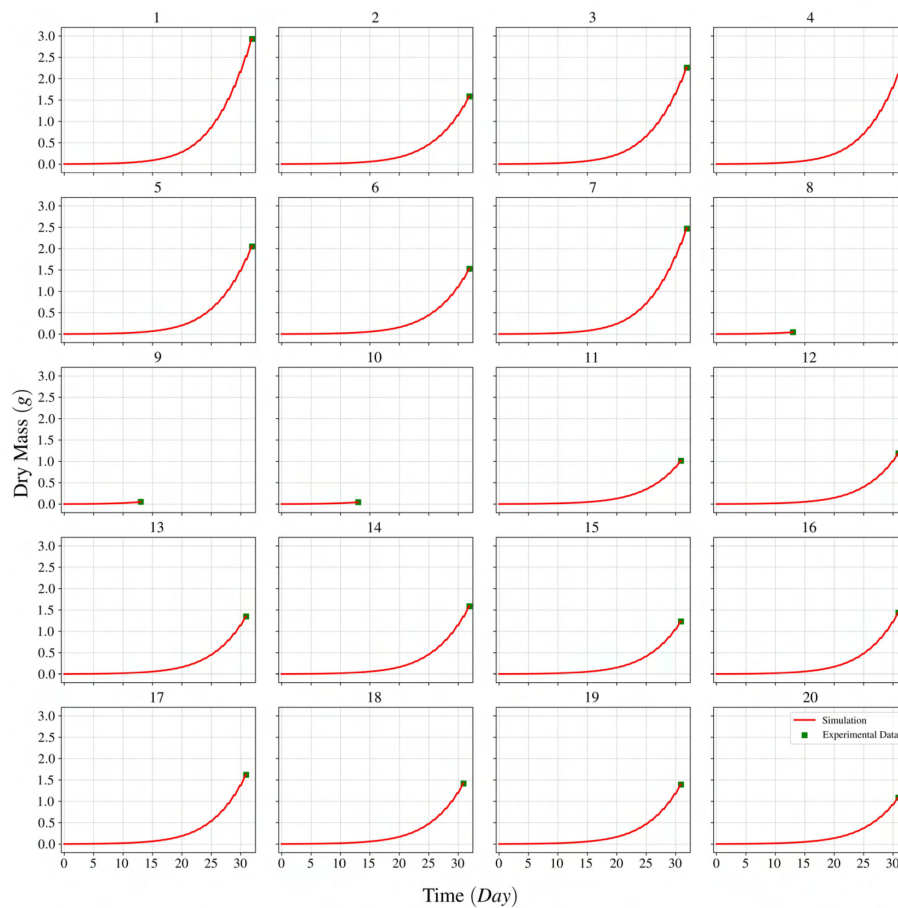


FIGURE 3 Comparison of lettuce dry mass for a dynamic growth model with experimental data under different spectral distributions and intensities. The label at the top of each graph represents a light treatment experiment according to Table 2.

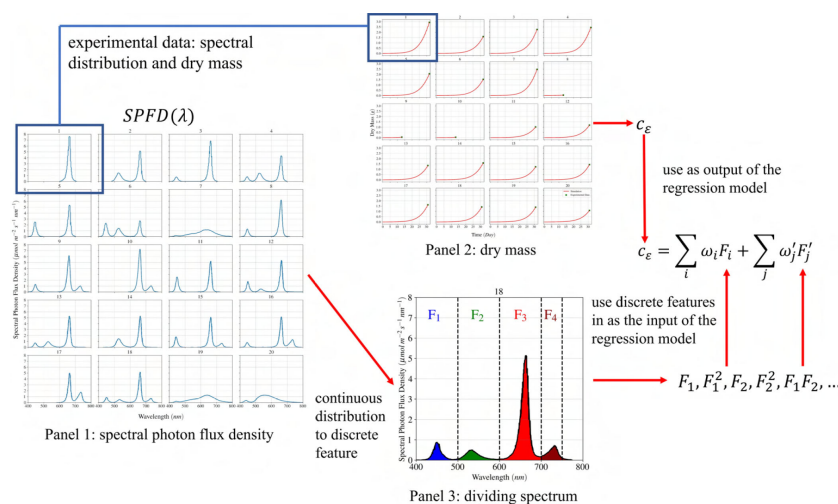


FIGURE 4 The development of the regression model for light-use efficiency based on spectral photon flux density distribution is shown. In Panel 1 (from Figure 1), λ is the wavelength (nm), and SPFD is the spectral photon flux density ($\mu \text{ mol m}^{-2} \text{ s}^{-1} \text{ nm}^{-1}$). The continuous SPFD is converted to four discrete features, as seen in Panel 3 (from Figure 5). Dry mass versus time is given in Panel 2 (from Figure 3), which is then converted to c_e . c_e , given on the right, is the light-use efficiency in the dynamic growth model of lettuce, and F_i and ω_i are the discrete input features based on the incoming spectrum and the corresponding weights, respectively. Finally, Terms with prime correspond to the interaction between different wavebands.

flux density for blue and green light is $20 \mu\text{mol m}^{-2} \text{s}^{-1}$, red light is $120 \mu\text{mol m}^{-2} \text{s}^{-1}$, and far-red light is $20 \mu\text{mol m}^{-2} \text{s}^{-1}$. The photon flux density for different wavebands represents the areas under the curve in Figure 5. Therefore, photon flux density ratios for the blue, green, and far-red wavebands are $\frac{20}{180}$, while for the red waveband is equal to $\frac{120}{180}$. Computation of the ratios for integrated spectral irradiation of the PAR+FR wavebands is different, since spectral irradiance is a measure of the energy carried by a photon. Therefore, integrated spectral irradiance (I) for a specific spectrum of PAR+FR is determined by calculating the energy for each wavelength through multiplication of wavelength energy and its number of photons and integrating those over the specific waveband. For “B20G20R120FR20” as an example, the integrated spectral irradiance for blue, green, red and far-red wavebands are 5.57 W m^{-2} , 4.57 W m^{-2} , 21.74 W m^{-2} , and 3.51 W m^{-2} , respectively. Based on the integrated spectral irradiance values computed for different wavebands, the integrated spectral irradiance intensity ratios for blue, green, red, and far-red are $\frac{5.57}{35.38}$, $\frac{4.57}{35.38}$, $\frac{21.74}{35.38}$, and $\frac{3.50}{35.38}$, or 15.7 : 12.9 : 61.4 : 9.9 respectively.

3.3 Exploratory data analysis

With the calculation of c_e from the experimental data, the next step is to create regression models to fit polynomial functions over a set of discrete variables and predict the light-use efficiency. As the aim of this study is to establish a framework that can estimate light-use efficiency as a function of incoming spectra, the input features include discrete parameters associated with either PFD or spectral irradiance

distribution. The first set of input variables consists of photon flux density ratios for blue, green, red, and far-red wavebands, which are calculated by integrating the photon flux density distribution shown in Figure 1. On the other hand, the second set of input parameters includes spectral irradiance ratios for the same wavebands, obtained by integrating over spectral irradiance distribution for various lighting distributions. Before investigating various regression models, the properties of the data used for regression are explored. Table 3 represents the mean, standard deviation (std), minimum (min), maximum (max), and percentile values for the investigated features (25% or first quartile, 50% or second quartile or median, 75% or third quartile), the input (F is the fraction ratio which is either based on photon flux density (PFD) or spectral irradiance integrated over wavelengths (I) whereas B, G, R, and FR represent blue, green, red, and far-red wavebands) and the output (c_e is the light-use efficiency) of the model. Figure 6 visualizes the 3D distribution for photon flux density and spectral irradiance fraction ratios for the studied dataset.

3.4 Predictive model including polynomial features for the quantum use efficiency coefficient (c_e)

Different polynomial models examined within the aim of this study have a form similar to

$$c_e = a_0 + \sum_{i=1}^4 \left[a_i F_i + b_i F_i^2 + c_i F_i^3 + d_i F_i^4 + \sum_{j=1}^4 (e_{ij} F_i F_j) \right]. \quad (11)$$

Not all of the models have every term presented in Equation (11), e.g., the regression model based on the first-order term is defined as $\bar{c}_e = a_0 + \sum_{i=1}^4 (a_i F_i)$. Regression models are defined in a way that includes up to 16 weight coefficients. Within the scope of this study, 22 distinctive terms are investigated, that are provided in Equation (11), and includes ratios (F_i , 4 terms representing each waveband), the square of ratios (F_i^2 , 4 terms), the cubic of ratios (F_i^3 , 4 terms), the quartic of ratios (F_i^4 , 4 terms), and interaction ratios ($F_i F_j$, 6 terms). Therefore, studied regression models within the scope of this study are comprised of linear models with 4 (includes F_i terms), 8, 12, 14, and 16 (includes F_i , F_i^2 , F_i^3 , F_i^4 terms) weight coefficients. Initially, a regression model with four terms corresponding to the first-order terms was developed and investigated as a baseline model. Higher-order polynomials were then explored to improve the poor performance ($R^2 < 0.45\%$) of this linear model. Polynomials with 8 (combinations of linear, second-order, and interaction terms), 12 (same with third-order terms), 14 (same as 12 with two more terms), and 16 (including fourth-order terms) were considered. It should be noted that all 22 terms (all combinations through fourth-order) were not used because of the small size of the dataset. Results from these model choices will be given in the following section.

Since the dataset is comprised of 20 observations of c_e for different spectral distributions and intensities (Figure 1), increasing the number of coefficients beyond the suggested limit would result in an overfitted model. In other words, this would lead to a model capable of accurate prediction for the studied data; however,

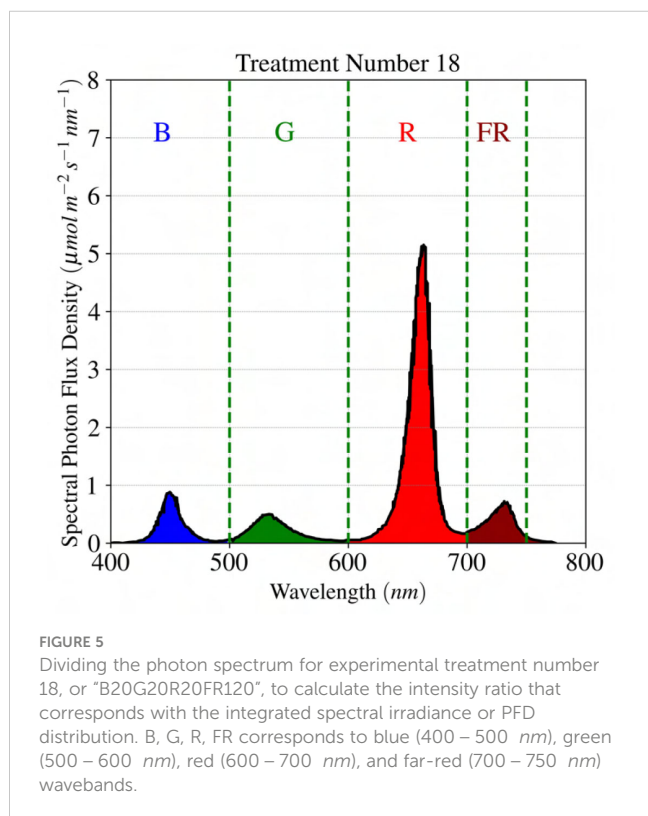


FIGURE 5
Dividing the photon spectrum for experimental treatment number 18, or “B20G20R120FR20”, to calculate the intensity ratio that corresponds with the integrated spectral irradiance or PFD distribution. B, G, R, FR corresponds to blue (400 – 500 nm), green (500 – 600 nm), red (600 – 700 nm), and far-red (700 – 750 nm) wavebands.

TABLE 3 Statistical information for ratios based on integrated spectral and photon flux density distributions, and c_e .

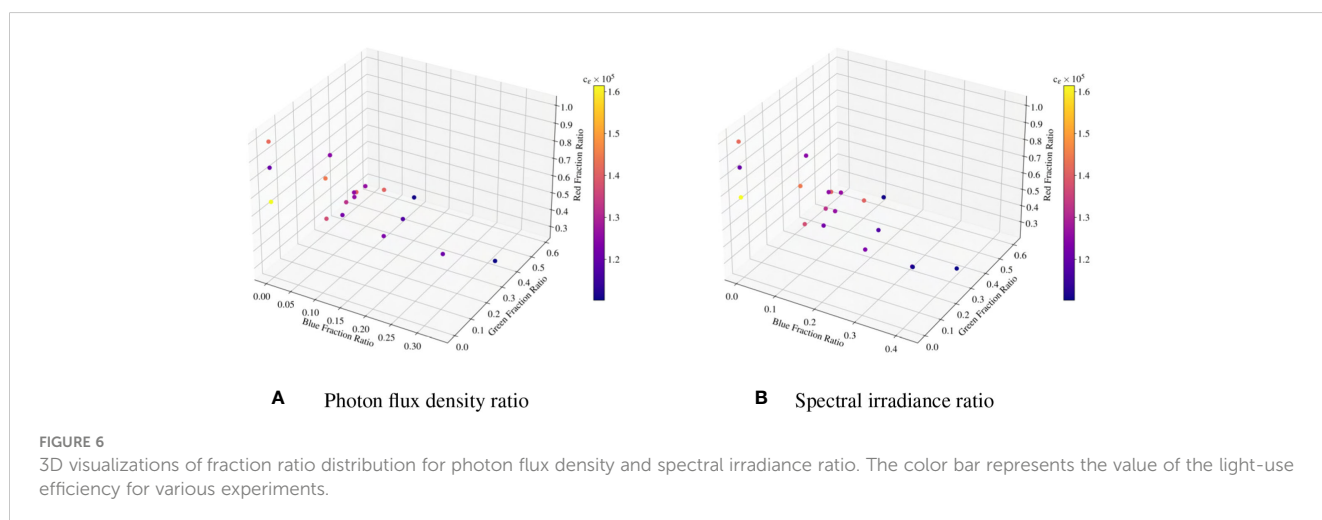
| Variable | Mean | Std | Min | 25% | 50% | 75% | Max |
|-------------------|-------|-------|-------|-------|-------|-------|-------|
| $F_{B, PFD}$ | 0.127 | 0.112 | 0 | 0.049 | 0.111 | 0.181 | 0.333 |
| $F_{G, PFD}$ | 0.147 | 0.178 | 0 | 0 | 0.056 | 0.309 | 0.588 |
| $F_{R, PFD}$ | 0.660 | 0.167 | 0.284 | 0.639 | 0.667 | 0.679 | 1 |
| $F_{FR, PFD}$ | 0.066 | 0.093 | 0 | 0 | 0 | 0.111 | 0.333 |
| $F_{B, I}$ | 0.168 | 0.142 | 0 | 0.065 | 0.155 | 0.250 | 0.421 |
| $F_{G, I}$ | 0.158 | 0.187 | 0 | 0 | 0.063 | 0.322 | 0.605 |
| $F_{R, I}$ | 0.614 | 0.174 | 0.254 | 0.562 | 0.612 | 0.665 | 1 |
| $F_{FR, I}$ | 0.060 | 0.088 | 0 | 0 | 0 | 0.097 | 0.321 |
| $c_e \times 10^5$ | 1.30 | 0.142 | 0.110 | 1.21 | 1.26 | 1.40 | 1.61 |

Mean, std, min, and max represent the average, standard variation, minimum, and maximum values in the dataset. 25 %, 50 %, and 75 % represent numerical values for the first quartile, median, and third quartile (based on the assumption that data is sorted in ascending order). F is the fraction ratio which is either based on photon flux density (PFD) or spectral irradiance integrated over wavelengths (I), whereas B, G, R, and FR represent blue, green, red, and far-red wavebands, and c_e is the light-use efficiency.

evaluating the performance of the model against new data would decrease the fidelity of the model. The term “nonlinear” in this section refers to polynomial models based on Equation (11) in which $e_{ij} \neq 0$ for every i and j value. The regression models which estimate the impact of incoming light spectrum on the light-use efficiency are classified into the following categories: 1) models based on PFD ratios that disregard the impact of overall photon flux density; 2) models based on integrated spectral irradiance ratios that disregard the impact of overall value; 3) models based on PFD ratios, that considered the impact of overall photon flux density; 4) models based on integrated spectral irradiance ratios that considered the impact of overall value. Figures 7A–D, represent the accuracy of categories 1, 2, 3, and 4 using R^2 metric.

To prevent overfitting and ensure the model’s applicability to new data, a validation study is conducted on the dataset. This

involves reserving a portion of the data for testing, which is not used during the model training phase. The testing data is used to evaluate the model’s performance on new and unseen data. The goal is to find a model that performs well on both training data and testing data, thereby preventing underfitting or overfitting issues. For instance, for a regression model with 16 weight coefficients, 3 samples are randomly chosen for testing purposes. The remaining 17 samples are used to train the regression model, and then its accuracy is evaluated on the 3 unseen samples. The selected regression model (and the corresponding weights for different terms) is the one that performs well on both training data (17 samples) and testing data (3 samples). This approach ensures that the model is not overfitting and can predict well on new data, making it useful for practical applications. In addition to using unseen data for the validation of the regression model, a



regularization penalty (L1 or L2 norm) is introduced into the regression model to decrease the variation caused by the complexity of the model.

3.5 Development and performance comparison of regression models

In this study, the Scikit-learn built-in function *LinearRegression* developed by Pedregosa et al. (2011) is used to build a model that minimizes the regularized residual sum of squares (R^2) as the criteria for the closest linear function to the actual data ¹. The R^2 score is calculated using $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$, in which y_i and \hat{y}_i represented the true and predicted value for the i^{th} sample and \bar{y} is the average of the actual values. In addition to R^2 criteria, the mean absolute percentage error (MAPE) is also calculated for different models and is defined as $MAPE = \frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)}$, where ϵ is an arbitrary non-zero small positive number to ensure that MAPE is defined. Figure 8 compares the actual value, $c_{e,simulation}$, versus the predicted value, $c_{e,prediction}$: a straight line would indicate a perfect prediction. In Figure 8A, only terms linear in the features F_i are included; in contrast, in Figure 8B, first-order and interaction terms in Equation (11) are used. It is seen that the accuracy increases with more features. Figure 8 provide qualitative insight into the accuracy, whereas R^2 and MAPE provide quantitative metrics, as shown in Figures 7, 9. According to Figures 7, 9, in spite of simplicity for the first-order and the combination of first and second-order models, these models have poor accuracy in predicting the c_e as a function of spectral distribution with R^2 score less than 0.75. This is addressed by considering several models created from the general expression in Equation (11): ten different combinations of features are shown, each as a horizontal bar. Not all combinations are shown in the Figures 7, 9–11; a representative set for a different number of terms is visualized, but one that includes the best performing model (16 terms).

Normalization of c_e to accommodate the impact of overall integrated spectral irradiance or photon flux density was examined and found to increase the fidelity of the regression model. Note the improvement for the particular case of a model that includes first-, second-, third-, and fourth-order terms (Fi+Se+Th+Fo bar in Figure 7), which we will establish as the best model below. For this case, using normalization resulted in a 16-feature regression model that has an average error of 2%, as compared to the unnormalized error of 15%.

In addition to R^2 and MAPE metrics, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are also computed for the studied models, which are the measures of the

model’s complexity. The value of AIC and BIC for models with a constant parameter is obtained using the following correlations (Seabold and Perktold, 2010): $AIC = -2 \times LL + \log(N) \times (k + 1)$, and $BIC = -2 \times LL + 2 \times (k + 1)$, whereas LL is the log of the likelihood function (how likely it is that the model predicted the actual values), N is the size of training data, and k is the number of features, respectively. It is worth mentioning that a lower value of AIC or BIC indicates that the fitted model is a better fit for the data, as it balances the goodness of fit and the complexity of the model (Baguley, 2018). Figures 10, 11 provide visual representations of the impact of the model’s complexity on AIC and BIC value. For the studied dataset, by increasing the accuracy of the model through the addition of new terms, the likelihood function also improves, which ultimately outweighs the negative penalty associated with the higher number of features (k); therefore, for the investigated models, the introduction of new features into the regression model would generally decrease AIC value. Figure 11 demonstrates that the BIC shows a different pattern, where regression models including nonlinear terms and only one of the first-, second-, or third-order terms are better fitted than those consisting of nonlinear terms and two of the first-, second-, or third-order terms.

It is important to note that while a lower AIC or BIC value suggests a better model, it does not necessarily mean that the model is the best possible fit for the data. In order to determine the best-fitted model, a comprehensive analysis of the performance of the studied models is conducted based on AIC, BIC, MAPE, and R^2 metrics. Based on this comparison, it is evident that the model, which incorporates first-, second-, third-, and fourth-order terms based on the integrated spectral irradiance ratio considering the impact of overall integrated spectral irradiance of light-use efficiency, performs better than the other models with higher accuracy, and relatively lower AIC and BIC values.

Equation (12) demonstrates the high fidelity regression model based on spectral irradiance distribution normalized with integrated spectral irradiance for R180 light treatment based on Table 2. As shown in Figure 12, predictions of the suggested model are in good accordance with numerical data (for the best performing model, the testing data include treatment numbers 7, 14, and 19.); however, additional data could prove useful in developing a more comprehensive model.

$$\begin{aligned}
 c_e/R_I = & -1.40 \times 10^{-4} + 1.82 \times 10^{-4} F_{B,I} + 2.06 \times 10^{-4} F_{B,I}^2 \\
 & - 7.71 \times 10^{-4} F_{B,I}^3 + 8.38 \times 10^{-4} F_{B,I}^4 + 2.32 \times 10^{-4} F_{G,I} \\
 & - 4.20 \times 10^{-4} F_{G,I}^2 + 1.53 \times 10^{-3} F_{G,I}^3 - 1.56 \times 10^{-3} F_{G,I}^4 \\
 & - 1.90 \times 10^{-4} F_{R,I} + 1.03 \times 10^{-3} F_{R,I}^2 - 1.10 \times 10^{-3} F_{R,I}^3 \\
 & + 4.13 \times 10^{-4} F_{R,I}^4 + 4.68 \times 10^{-4} F_{FR,I} - 5.25 \times 10^{-3} F_{FR,I}^2 \\
 & + 2.92 \times 10^{-2} F_{FR,I}^3 - 4.77 \times 10^{-2} F_{FR,I}^4
 \end{aligned}
 \tag{12}$$

Equation (12) is constrained by the following condition,

$$F_{B,I} + F_{G,I} + F_{R,I} + F_{FR,I} = 1
 \tag{13}$$

Now that the best-performing model is selected, the impact of adding higher-order terms to the regression model is investigated. The

1 In order to mitigate potential overfitting, the linear regression was also performed using both *ridge* and *LASSO* regularization, again using Scikit-learn built-in function. Figure 7 indicates the importance of high-order polynomial terms but is not regularized; however, results in other figures employ ridge regression.

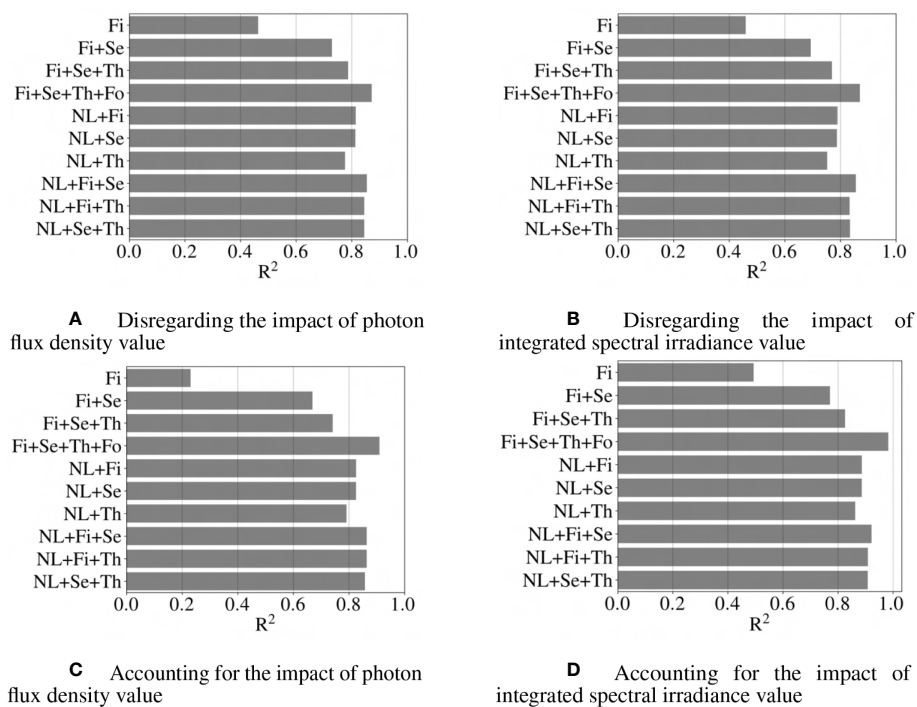


FIGURE 7
 Comparison of different models based on R^2 score criteria. Abbreviations within these figures are NL, nonlinear (includes 6 interaction terms between wavebands), Fi, first-order terms; Se, second-order terms; Th, third-order terms; Fo, fourth-order terms. R^2 score closer to 1 indicates a more accurate model.

following notation $[a_1, a_2, a_3, a_4]$ is used to evaluate the impact of higher-order term addition on first-order terms, in which a_1, a_2, a_3 , and a_4 represent weights for blue, green, red, far-red fraction ratios, respectively. The examined models included a first-order term model with weights of $[-1.24 \times 10^{-5}, -6.96 \times 10^{-6}, -3.60 \times 10^{-6}, -$

$2.18 \times 10^{-7}]$, a combination of first- and second-order terms model with $[6.95 \times 10^{-5}, 7.34 \times 10^{-5}, 8.02 \times 10^{-5}, 4.29 \times 10^{-5}]$, a combination of first- to third-order terms model with $[6.89 \times 10^{-6}, -2.41 \times 10^{-5}, 4.60 \times 10^{-5}, -1.60 \times 10^{-5}]$, and a combination of first- to fourth-order terms model with $[1.82 \times 10^{-4}, 2.32 \times 10^{-4},$

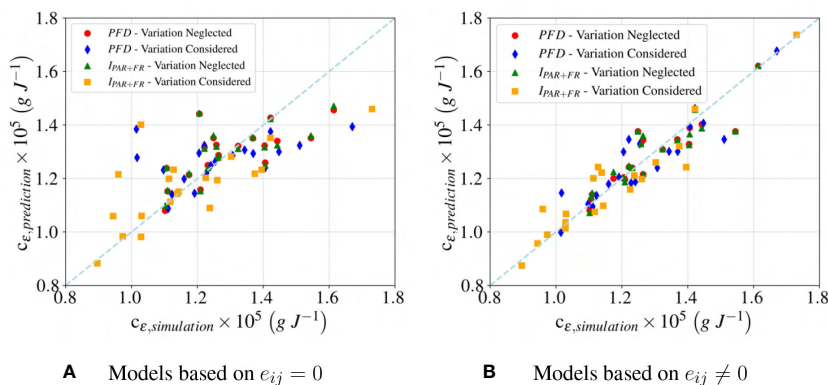


FIGURE 8
 Accuracy of first-order regression models with the assumption of (A) Neglecting nonlinearity ($e_{ij} = 0$), and (B) Considering nonlinearity ($e_{ij} \neq 0$). IPAR+FR refers to a regression model based on the integrated spectral irradiance ratio, while PFD represents a model based on the photon flux density ratio.

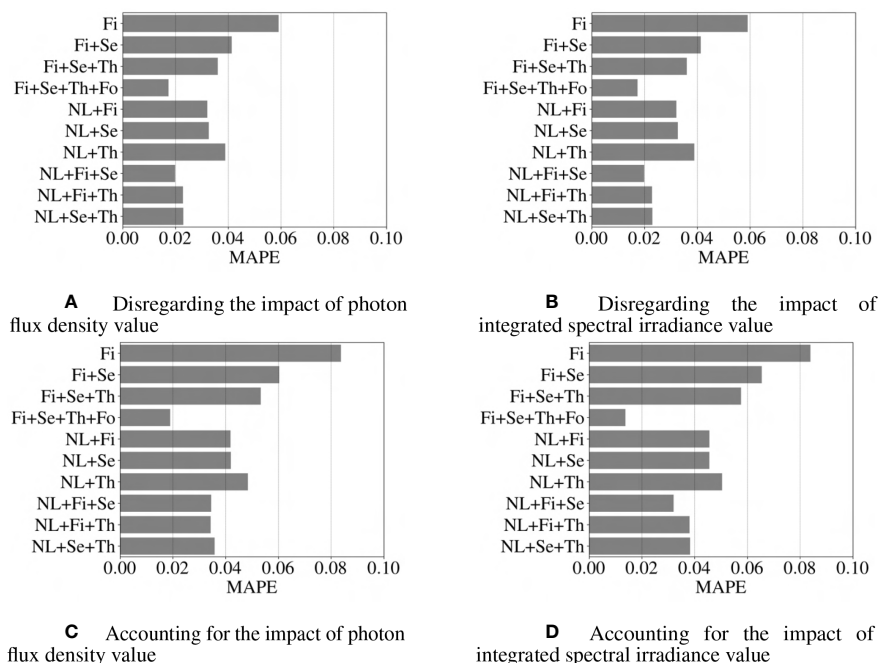


FIGURE 9

Comparison of different models based on the mean absolute percentage error (MAPE) score criteria. Abbreviations within these figures are NL, nonlinear (includes 6 interaction terms between wavebands), Fi, first-order terms; Se, second-order terms; Th, third-order terms; Fo, fourth-order terms. The lower value for MAPE score indicates a more accurate model.

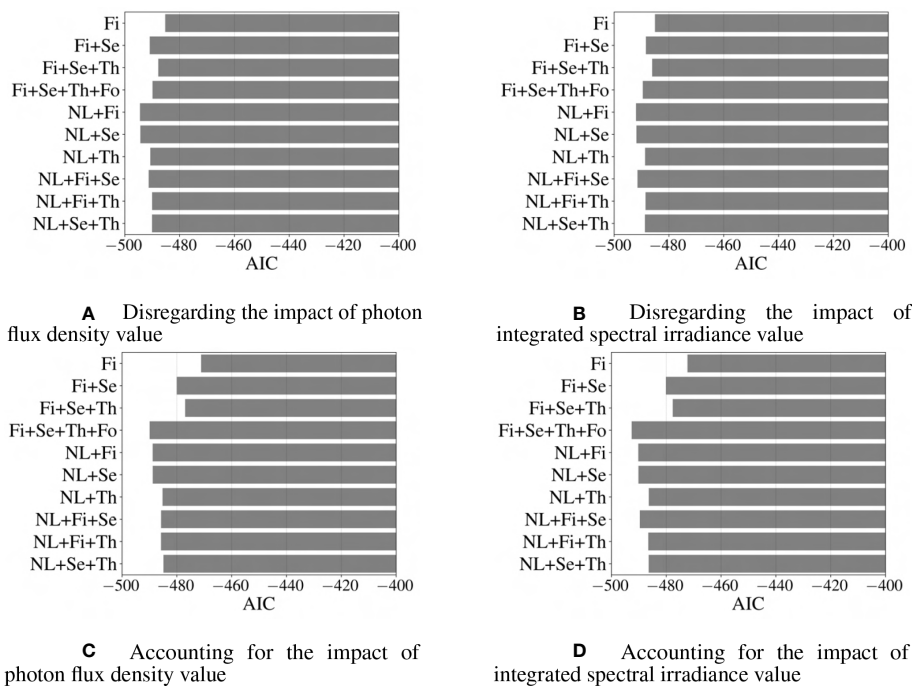


FIGURE 10

Comparison of different models based on AIC criteria. Abbreviations within these figures are NL, nonlinear (includes 6 interaction terms between wavebands), Fi, first-order terms; Se, second-order terms; Th, third-order terms; Fo, fourth-order terms. The lower the value for AIC, the better the fit of the model (Baguley, 2018).

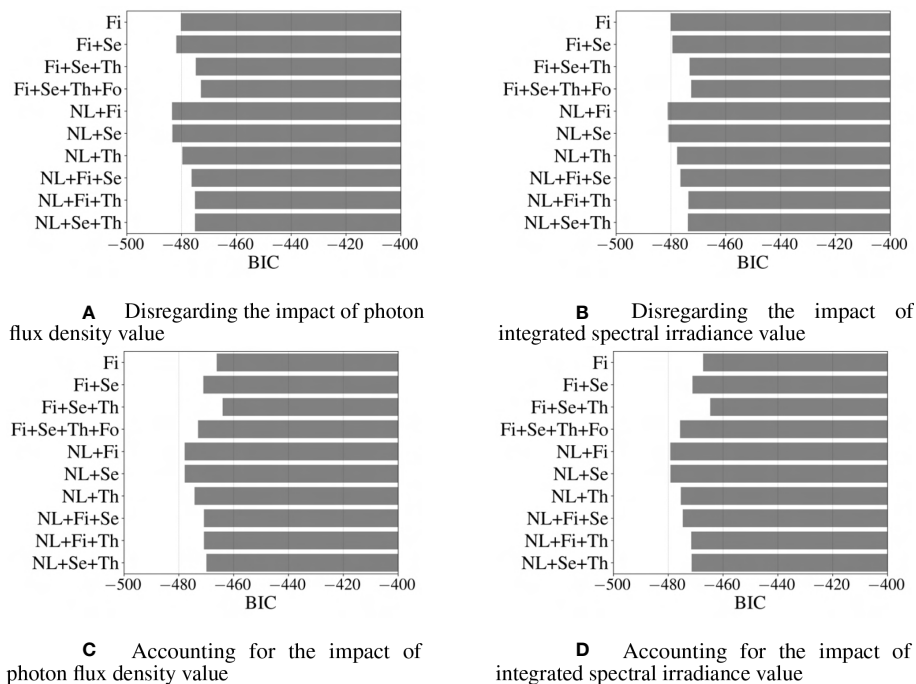


FIGURE 11 Comparison of different models based on BIC criteria. Abbreviations within these figures are NL, nonlinear (includes 6 interaction terms between wavebands), Fi, first-order terms; Se, second-order terms; Th, third-order terms; Fo, fourth-order terms. The lower the value for BIC, the better the fit of the model (Baguley, 2018).

-1.90×10^{-4} , 4.68×10^{-4}]. Comparison between these notations suggests that the addition of higher-order terms to the regression models significantly affects the weights of the first-order terms, highlighting the importance of exploring more complex models with higher-order terms for improved prediction accuracy.

3.6 Coupling the regression model with van henten dynamic growth model for lettuce

Using Equation (12) and integrated spectral irradiance fraction ratio for the investigated wavebands of the different light treatment

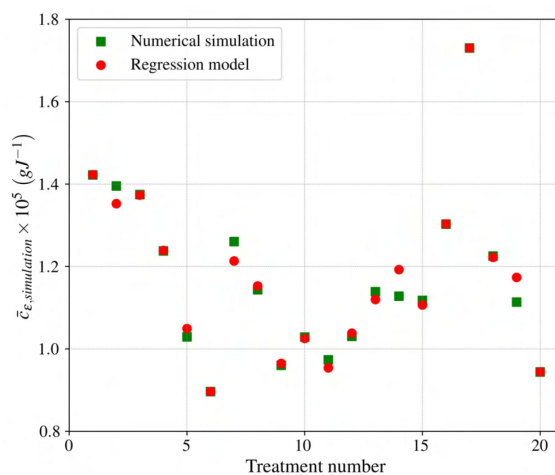


FIGURE 12 Comparison of c_e for the suggested model with numerical simulation for different light treatments.

cases, the Van Henten (1994) dynamic growth model is utilized to compare the accuracy of the modified growth model that accounts for spectral distribution and intensity. Figure 13 compares the difference between experimental dry mass for lettuce with dry mass prediction of the modified growth model.

Comparing the numerical error of the lettuce growth model using the suggested value for light-use efficiency (labeled constant in Figure 13), with that of the lettuce growth model with the proposed regression model (Equation (11); labeled regression in Figure 13), it is inferred that the suggested regression model of the c_e improves the accuracy of Van Henten (1994) growth model and adequately considers the impact of spectral distribution on plant growth.

3.7 Cross-validation of the proposed light-use efficiency model with novel data

In the previous section, the performance of the proposed mathematical model is investigated through integration with the Van Henten (1994) dynamic growth model. To further investigate the accuracy of the light-use efficiency model based on the spectral irradiance ratios of various wavebands, data from Both et al. (1994) on lettuce cultivated in a controlled greenhouse environment is utilized. From October 1992 to March 1993, six controlled light treatments (without the use of supplemental lighting) were conducted for lettuce grown hydroponically. For these treatments, during the first 11 days, the temperature and CO₂ concentration were maintained at 25 °C and 350 mL L⁻¹, respectively. After day 11, the temperature was set to 24 °C between 7 am and 5 pm and 18.8 °C for the rest of the day, while CO₂ was enriched to 1000 mL L⁻¹. Table 4 represents data used for cross-validation of the suggested regression-based light-use efficiency model. Spectral irradiance intensity and distribution ratios are approximated using reported daily light integral for the different treatments and predicted solar irradiation by Tobiska et al. (2000). Figure 14 displays a comparison of the spectral photon flux density distribution with respect to wavelength for the investigated LED spectrums and natural lighting.

According to Figure 15, using the proposed regression-based light-use efficiency model, lettuce dry mass predictions were in close agreement (mostly within 2-3 percent error) with the experimental data.

4 Discussion and conclusion

The aim of this study is to predict the impact of incoming light spectral distribution and its intensity on lettuce growth. For this purpose, a dynamic model of plant growth for lettuce provided by Van Henten (1994) is modified. An ODE solver is developed to simulate the dynamic behavior of lettuce from the seedling stage to maturity. It is assumed that the spectral distribution of light and its intensity affect the model through a coefficient c_e , which accounts for energy provided by photons for a reduction of one molecule of CO₂. Using data for lettuce cultivated under 20 different indoor lighting treatments, the ODE solver calculated c_e for different cases. Several models are fitted using spectral distribution ratios for 4 light wavebands: blue (400 – 500 nm), green (500 – 600 nm), red (600 – 700 nm), and far-red (700 – 750 nm) as input data and the obtained c_e as the sole output. To determine the algebraic structure of the model with the highest accuracy, a variety of regression models with varying numbers of features, from 4 (ratios for the blue, green, red, and far-red wavebands) to 16 (first, second, third, and fourth-order values for these ratios) are investigated. The combination of first to fourth-order terms that had the highest accuracy (98 %) was a regression model based on integrated spectral irradiance distribution (in which the predicted c_e was based on normalized overall spectral irradiance). In order to obtain coefficients for different terms in the regression model, 17 of the 20 experimental data were utilized, while the rest prevented the overfitting of the regression model. To further evaluate the accuracy of the regression model, 21 experimental data for three replications of indoor-cultivated lettuce were used (Both et al., 1994) and are presented in Figure 15.

The impact of incoming spectral distribution on lettuce plant growth is investigated. By considering two constrained scenarios, it is possible to visualize the impact of varying spectral distribution on

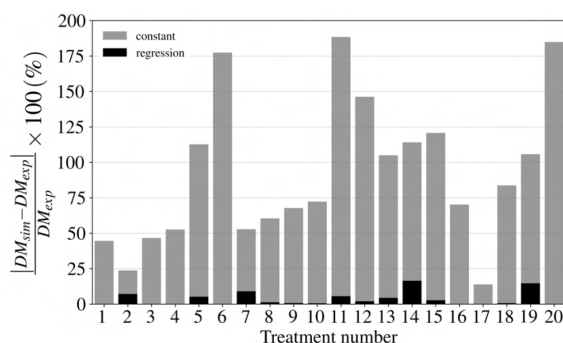


FIGURE 13 Computed numerical error is significantly reduced based on c_e value using Equation (12) (regression label, black bar) compared with error using the suggested constant value for v_e by Van Henten (1994) (constant label, light gray bar). DM_{sim} , and DM_{exp} are lettuce dry mass for numerical simulation and experimental study in g, respectively. As shown in the figure, coupling the regression model with the dynamic growth model improved the accuracy of prediction for lettuce cultivated under different spectral distributions.

TABLE 4 Experimental dry mass data for greenhouse cultivated hydroponic lettuce grown under controlled light treatments obtained by Both et al. (1994).

| Experimental period | Daily light integral (mol m ⁻² day ⁻¹) | Dry Mass | | | | | | |
|---------------------|---|----------|--------|--------|--------|--------|--------|--------|
| | | Day 14 | Day 18 | Day 21 | Day 25 | Day 28 | Day 32 | Day 35 |
| November 1992 | 6.2 | 0.064 | 0.153 | 0.3 | 0.76 | 1.01 | 1.735 | 2.46 |
| January 1993 | 4.7 | 0.063 | 0.141 | 0.2 | 0.44 | 0.84 | 1.2 | 1.87 |
| February 1993 | 10.5 | 0.085 | 0.288 | 0.53 | 1.21 | 1.98 | 3.15 | 4.81 |

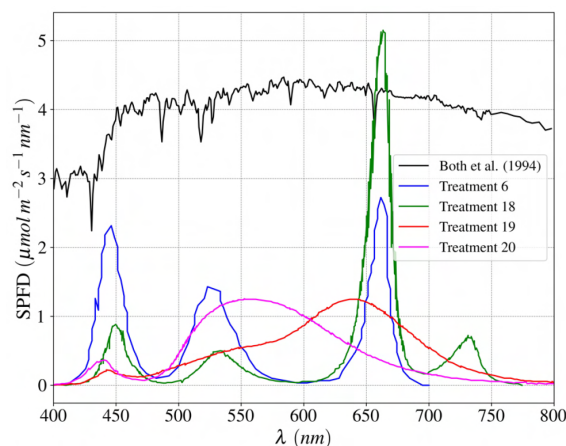


FIGURE 14 Comparison of spectral photon flux density distribution for natural light (Both et al., 1994) and investigated LEDs spectrum.

light-use efficiency. In these scenarios, the spectral irradiance integrated over wavelengths (I_{PAR+FR}) remains unchanged while one of the wavebands is eliminated from the spectra. In the first scenario, it is assumed that the far-red waveband is missing from the light spectra and only contains the traditionally defined PAR waveband. Contrary to the first scenario, in the second one, the green waveband is replaced with far-red; thus, the incoming spectrum is composed of blue, red, and far-red wavebands. Figures 16A, B demonstrate how light-use efficiency varies in scenarios one and two, respectively. Based on these figures, the red waveband promotes and the blue waveband inhibits biomass

accumulation in lettuce by increasing and decreasing light-use efficiency, respectively. The impact of the blue waveband on plant growth is enhanced by the presence of the green waveband. On the other hand, the addition of the far-red waveband to the light spectrum mitigates the impact of the blue/green waveband. Therefore, maximum lettuce biomass accumulation can be achieved using the outcomes of these scenarios by emphasizing the red and far-red wavebands and avoiding a high spectral irradiance ratio of the blue and green bands. However, this model ignores important quality considerations such as leaf color, texture, nutritional content, and post-harvest longevity.

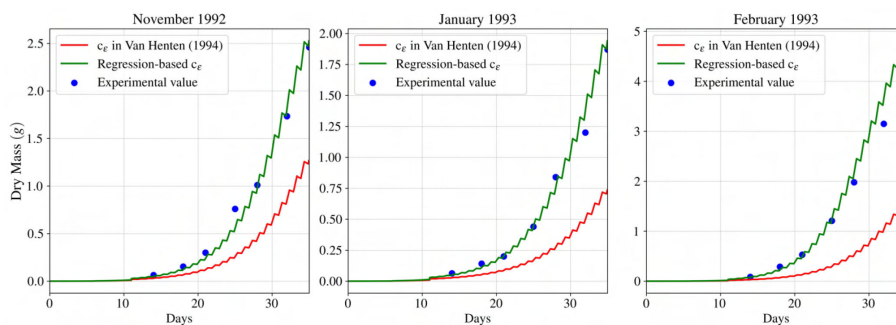


FIGURE 15 Comparison of dynamic growth model accuracy using the suggested value for c_e by Van Henten (1994), and obtained value using the proposed regression-based light-use efficiency model with experimental data from Both et al. (1994) for periods of November 1992, and January and February of 1993. The regression-based model is capable of approximating dry mass for greenhouse cultivated lettuce.

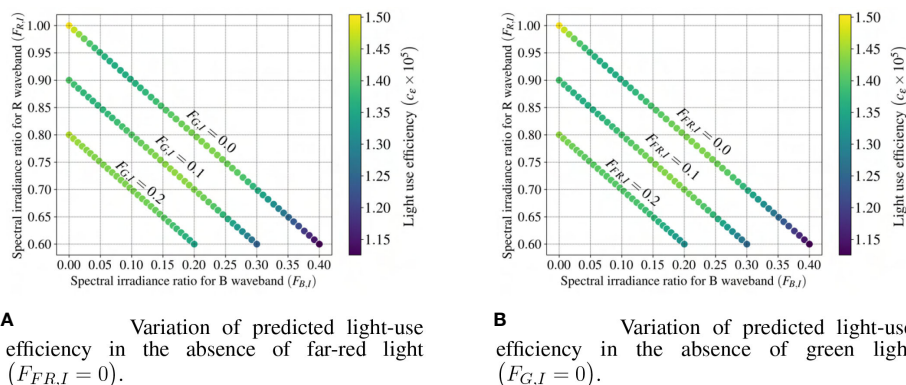


FIGURE 16

Impact of the incoming light spectral distribution on the light-use efficiency coefficient (c_e) considering the same spectral irradiance integrated over wavelength (IPAR+F R remains unchanged). B and R refer to blue and red. The constraint of Equation (13) enforces a linear relationship between $F_{B,I}$ and $F_{R,I}$ within each scenarios, and ensures a clear visualization of the effect of the spectral distribution on c_e .

The model presented in Equation (12) provides a simplified framework to evaluate the impact of spectral distribution on lettuce plant growth. This contrasts with a model for tomato growth (Dieleman et al., 2019), in which the 3D model needs to be solved in order to investigate the effect of light quality. Moreover, the application of this model to the cultivation of lettuce can increase biomass accumulation during the plant growth cycle. This model can also be used to optimize light conditions, allowing for more efficient use of energy and resources.

Although the regression model predictions are in good accordance with the solver predictions, additional experimental data with a focus on the impact of light spectrum on lettuce plant morphology will likely create a more comprehensive model with higher fidelity. Furthermore, adding higher-order terms to the regression model resulted in a decrease in the weights of the first-order terms, which in turn suggests the necessity of investigating regression models with higher complexity on more comprehensive data for developing an accurate model of light-use efficiency. Since in the studied dataset, the number of samples is limited to 20 experiments, it is not feasible to assess the performance of more complex regression models. In addition, interactions likely exist between light intensity and the photon spectrum, and additional data are needed to test and improve the model's performance. Specifically, morphological acclimation, such as total leaf area, canopy area, number of leaves, leaf pigmentation, and chlorophyll concentration, can affect the photosynthetic rate and light interception and would ideally be parameterized in future growth models. Finally, this technique has the potential to be applied to other horticultural crops, particularly leafy vegetable crops, to incorporate the impact of spectral distribution on biomass accumulation and crop yield.

Data Availability Statement

The data analyzed in this study is subject to the following licenses/restrictions: Data was extracted from published articles cited in the references. Requests to access these datasets should be directed to benard@msu.edu.

Author contributions

MA: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing. XT: Methodology, Software, Review, and editing. ES: Investigation, and Review. ER: Review, Editing, Funding acquisition. JK: Review and editing, Funding acquisition. MM: Review and editing. AB: Supervision, Review and editing, Conceptualization, Project administration, and Funding acquisition. All authors contributed to the article and approved the submitted version.

Acknowledgments

We gratefully acknowledge support from the National Institute of Food and Agriculture from grant number 2018-67003-27407 titled "INFEWS/T3: Advanced Energy Efficient Greenhouse Systems Employing Spectral Splitting and Solar Water Purification". We would like to express our gratitude toward Dr. Qingwu (William) Meng, Assistant Professor at the Department of Plant and Soil Sciences, at the University of Delaware, who provided us with environmental data for his studies on the impact of different wavebands on lettuce plant growth.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Acock, B., Charles-Edwards, D., Fitter, D., Hand, D., Ludwig, L., Warren Wilson, J., et al. (1978). The contribution of leaves from different levels within a tomato crop to canopy net photosynthesis: an experimental examination of two canopy models. *J. Exp. Bot.* 29, 815–827. doi: 10.1093/jxb/29.4.815
- Baguley, T. (2018). *Serious stat: a guide to advanced statistics for the behavioral sciences* (Bloomsbury publishing).
- Both, A., Albright, L., Langhans, R., Reiser, R., and Vinzant, B. (1994). “Hydroponic lettuce production influenced by integrated supplemental light levels in a controlled environment agriculture facility: experimental results,” in *Acta hortic.*, vol. 45–52. (Noordwijkerhout, Netherlands: International Society for Horticultural Science). doi: 10.17660/ActaHortic.1997.418.5
- Dieleman, J. A., De Visser, P. H., Meinen, E., Grit, J. G., and Dueck, T. A. (2019). Integrating morphological and physiological responses of tomato plants to light quality to the crop level by 3d modeling. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00839
- Goudriaan, J., and Monteith, J. (1990). A mathematical function for crop growth based on light interception and leaf area expansion. *Ann. Bot.* 66, 695–701. doi: 10.1093/oxfordjournals.aob.a088084
- Goudriaan, J., Van Laar, H., Keulen, H. V., and Louwse, W. (1985). “Photosynthesis, CO₂ and plant production,” in *Wheat growth and modelling* (Springer), 107–122. doi: 10.1007/978-1-4899-3665-3
- He, R., Zhang, Y., Song, S., Su, W., Hao, Y., and Liu, H. (2021). Uv-a and fr irradiation improves growth and nutritional properties of lettuce grown in an artificial light plant factory. *Food Chem.* 345. doi: 10.3389/fpls.2021.667407
- Hebert, D., Boonekamp, J., Parrish, C. H., Ramasamy, K., Makarov, N. S., Castañeda, C., et al. (2022). Luminescent quantum dot films improve light use efficiency and crop quality in greenhouse horticulture. *Front. Chem.* 10, 1345. doi: 10.3389/fchem.2022.988227
- Hogewoning, S. W., Wientjes, E., Douwstra, P., Trouwborst, G., Van Ieperen, W., Croce, R., et al. (2012). Photosynthetic quantum yield dynamics: from photosystems to leaves. *Plant Cell* 24, 1921–1935. doi: 10.1105/tpc.112.097972
- Jones, J. W., Dayan, E., Allen, L., Van Keulen, H., and Challa, H. (1991). A dynamic tomato growth and yield model (tomgro). *Trans. ASABE* 34, 663–667. doi: 10.13031/2013.31715
- Kelly, N., Choe, D., Meng, Q., and Runkle, E. S. (2020). Promotion of lettuce growth under an increasing daily light integral depends on the combination of the photosynthetic photon flux density and photoperiod. *Sci. Hortic.* 272, 109565. doi: 10.1016/j.scienta.2020.109565
- Kong, Y., and Nemali, K. (2021). Blue and far-red light affect area and number of individual leaves to influence vegetative growth and pigment synthesis in lettuce. *Front. Plant Sci.* 12, 667407. doi: 10.1016/j.fplant.2020.128727
- Lorenz, H., and Wiebe, H. (1980). Effect of temperature on photosynthesis of lettuce adapted to different light and temperature conditions. *Sci. Hortic.* 13, 115–123. doi: 10.1016/0304-4238(80)90075-8
- Meng, Q., Boldt, J., and Runkle, E. S. (2020). Blue radiation interacts with green radiation to influence growth and predominantly controls quality attributes of lettuce. *J. Am. Soc. Hortic.* 145, 75–87. doi: 10.21273/JASHS04759-19
- Meng, Q., Kelly, N., and Runkle, E. S. (2019). Substituting green or far-red radiation for blue radiation induces shade avoidance and promotes growth in lettuce and kale. *Environ. Exp. Bot.* 162, 383–391. doi: 10.1016/j.envexpbot.2019.03.016
- Meng, Q., and Runkle, E. S. (2019). Far-red radiation interacts with relative and absolute blue and red photon flux densities to regulate growth, morphology, and pigmentation of lettuce and basil seedlings. *Sci. Hortic.* 255, 269–280. doi: 10.1016/j.scienta.2019.05.030
- Ooms, M. D., Dinh, C. T., Sargent, E. H., and Sinton, D. (2016). Photon management for augmented photosynthesis. *Nat. Commun.* 7, 1–13. doi: 10.1038/ncomms12699
- Park, Y., and Runkle, E. S. (2018a). Far-red radiation and photosynthetic photon flux density independently regulate seedling growth but interactively regulate flowering. *Environ. Exp. Bot.* 155, 206–216. doi: 10.1016/j.envexpbot.2018.06.033
- Park, Y., and Runkle, E. S. (2018b). Spectral effects of light-emitting diodes on plant growth, visual color quality, and photosynthetic photon efficacy: white versus blue plus red radiation. *PLoS One* 13, e0202386. doi: 10.1371/journal.pone.0202386
- Park, Y., and Runkle, E. S. (2019). Blue radiation attenuates the effects of the red to far-red ratio on extension growth but not on flowering. *Environ. Exp. Bot.* 168, 103871. doi: 10.1016/j.envexpbot.2019.103871
- Paz, M., Fisher, P. R., and Gómez, C. (2019). Minimum light requirements for indoor gardening of lettuce. *Urban Agric. Reg. Food Syst.* 4, 1–10. doi: 10.2134/urbanag2019.03.0001
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Sakamoto, M., and Suzuki, T. (2015). Effect of root-zone temperature on growth and quality of hydroponically grown red leaf lettuce (*Lactuca sativa* L. cv. red wave). *Am. J. Plant Sci.* 6, 2350. doi: 10.4236/ajps.2015.614238
- Seabold, S., and Perktold, J. (2010). Statsmodels: econometric and statistical modeling with python. In *Proc. 9th Python Sci. Conf. (Austin TX)* 57, 10–25080. doi: 10.25080/Majora-92bf1922-011
- Shen, L., Lou, R., Park, Y., Guo, Y., Stallknecht, E. J., Xiao, Y., et al. (2021). Increasing greenhouse production by spectral-shifting and unidirectional light-extracting photonics. *Nat. Food* 2, 434–441. doi: 10.1038/s43016-021-00307-8
- Stanghellini, C. (1987). Transpiration of greenhouse crops: an aid to climate management (Wageningen university and research).
- Sweeney, D. (1981). Modelling the growth of winter lettuce. *Math. Plant Physiol.* 217–229.
- Tobiska, W. K., Woods, T., Eparvier, F., Viereck, R., Floyd, L., Bouwer, D., et al. (2000). The solar2000 empirical solar irradiance model and forecast tool. *J. Atmos. Sol. Terr. Phys.* 62, 1233–1250. doi: 10.1016/S1364-6826(00)00070-5
- Van Henten, E. (1994). Validation of a dynamic lettuce growth model for greenhouse climate control. *Agric. Syst.* 45, 55–72. doi: 10.1016/S0308-521X(94)90280-1
- Van Henten, E., and Van Straten, G. (1994). Sensitivity analysis of a dynamic growth model of lettuce. *J. Agric. Eng. Res.* 59, 19–31. doi: 10.1006/jaer.1994.1061
- Van Holsteijn, H. M. (1981). Growth and photosynthesis of lettuce (Wageningen university and research).

Glossary

| | |
|----------------|---|
| C_{CO_2} | Carbon dioxide concentration ($mL L^{-1}$) |
| DLI | Daily light integral ($mol m^{-2} day^{-1}$) |
| F | Intensity ratio corresponding to specific wavelengths in the form of $(\frac{PFD_{specific}}{PFD_{overall}})$ or $(\frac{I_{specific}}{I_{overall}})$ |
| f | Mass rate of change ($g m^{-2} s^{-1}$) |
| g | Conductance ($m s^{-1}$) |
| I | Spectral irradiance integrated over wavelengths ($W m^{-2}$) |
| $PPFD$ | Photon flux density ($\mu mol m^{-2} s^{-1}$) |
| R_t | Ratio of spectral irradiance integrated over PAR+FR wavebands to a spectral irradiance reference value (R180 is considered as the reference) $(\frac{I_{PAR+FR}}{I_{PAR+FR,R180}})$ |
| r_{gr} | Growth rate of structural material ($g m^{-2} s^{-1}$) |
| $SPFD$ | Spectral photon flux density ($\mu mol m^{-2} s^{-1} nm^{-1}$) |
| T | Temperature ($^{\circ}C$) |
| X | Dry Mass ($g m^{-2}$) |
| Subscripts: | |
| B | Blue light waveband (400 – 500 nm) |
| bnd | Boundary layer |
| car | Carboxylation |
| FR | Far-red light waveband (700 – 750 nm) |
| G | Green light waveband (500 – 600 nm) |
| K | Extinction coefficient |
| lar | Leaf area ratio |
| max | Saturation rate ($g m^{-2} s^{-1}$) |
| $nsdm$ | Non-structural dry mass ($g m^{-2}$) |
| PAR | Photosynthetically active radiation (400 – 700 nm) |
| $PPFD$ | Photosynthetic photon flux density ($\mu mol m^{-2} s^{-1}$) |
| $phot$ | Carbon dioxide photosynthesis ($gm^{-2} s^{-1}$) |
| Q_{10} | Q_{10} factor |
| R | Red light waveband (600 – 700 nm) |
| $resp$ | Maintenance respiration ($gm^{-2} s^{-1}$) |
| rt | Root |
| sdm | Structural dry mass ($g m^{-2}$) |
| sht | Shoot |
| stm | Stomata |
| Greek Letters: | |
| α | Conversion of assimilated CO_2 |
| β | Yield factor |

(Continued)

Continued

| | |
|------------|---|
| γ | Growth rate coefficient |
| Γ | CO_2 compensation point ($ml L^{-1}$) |
| ϵ | Light-use efficiency ($g J^{-1}$) |
| τ | Ratio of root dry mass to total dry mass |
| ω | Density of CO_2 ($g m^{-3}$) |