Smart Agricultural Technology 13 (2026) 101666

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

An extended growth model for *lactuca sativa* under controlled water deficit conditions

Davide Marino [a], Alessandro Antona [b], Arianna Catenacci [c], Gianni Ferretti [a,*]

[a] Politecnico di Milano, Department of Electronics, Information and Bioengineering, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy
[b] Agricola Moderna, Via Sandro Pertini, 26, 20066, Melzo, Italy
[c] Politecnico di Milano, Department of Civil and Environmental Engineering, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy

ARTICLE INFO

Keywords:
Lettuce growth modeling
Indoor farming
Dynamic crop model
Water stress
Substrate water content
Controlled environment agriculture.

ABSTRACT

Indoor vertical farming offers sustainable food production with higher yields and lower resource use through complete environmental control and advanced management systems. However, these systems require accurate models, based on either data-driven approaches or fundamental principles. This paper focuses on the extension of a proposed dynamic growth model for lettuce (*Lactuca sativa*), one of the most commonly cultivated crops in indoor farms due to its short growth cycle. The extension integrates the substrate water content as a dynamic state variable to simulate the effects of water stress on biomass accumulation and links it to the original model through a water stress coefficient dependent on these dynamics. Validation was carried out using experimental data from a commercial indoor farm, where lettuce was grown on two substrates (peat and wood fiber) under two different irrigation regimes. Following sensitivity and collinearity analyses, the model was calibrated with respect to light use efficiency and substrate water content at wilting point parameters. The extended model demonstrated high accuracy in predicting dry weight production under all experimental conditions. These results highlight the model's potential as a decision-support tool for optimizing irrigation strategies and substrate selection in indoor farming systems.

1. Introduction

Agriculture in controlled environments has become an increasingly attractive option in recent years due, to its high food production, low environmental impact and reduced resource consumption. Controlled environments enable the regulation of environmental factors and a year-round cultivation, while optimizing water, energy and space.

Two main types of controlled environments are used in modern agriculture: greenhouses and indoor vertical farms, the main differences between the two are the use of natural or artificial inputs and the possible influence of outside climate conditions. Greenhouses are semi-controlled environments that rely partly on natural light, and their internal climate is still influenced by external climate conditions. On the other hand, indoor farming is fully isolated from the outside and uses artificial inputs sources, with lighting usually provided by LEDs.

This total environmental control makes indoor farming an ideal platform for implementing sophisticated agricultural management techniques. In this respect, when selecting methods for managing environmental conditions, there are two primary control approaches available: data-driven strategies and model-based methodologies. In this study, we employ a model-based control framework to characterize and optimize crop-environment interactions. This approach is essential for clarifying the objectives and scope of the present work, which focuses on developing a transparent and interpretable modeling strategy rather than a purely predictive one.

The data-driven control approach utilizes a large volume of sensor and actuator data to optimize environmental conditions, without relying on detailed plant models. By continuously collecting data, such as temperature, humidity, CO_2 concentration, light intensity, and substrate water content, Machine Learning (ML) and statistical models like Neural Networks, Random Forests, or Reinforcement Learning agents can learn the complex, nonlinear and dynamic relationships between environmental inputs and plant characteristics, such as growth rate, biomass, yield, or quality. For example, [1] employed ML and Deep Learning techniques to predict yield and plant growth variation in tomato yield forecasting and *Ficus benjamina* stem growth, in controlled greenhouse environments. While data-driven methods offer flexibility and adaptability, they face limitations, such as the need for large high-quality datasets that are not always available or can only be acquired with destructive sampling. Moreover, the black-box nature of most ML models can make it diffi-

* Corresponding author at: Politecnico di Milano, Cremona Campus, Via Stefano Leonida Bissolati, 34, 26100, Cremona, Italy.
E-mail address: gianni.ferretti@polimi.it (G. Ferretti).

https://doi.org/10.1016/j.atech.2025.101666
Received 10 October 2025; Received in revised form 19 November 2025; Accepted 21 November 2025
Available online 24 November 2025

cult to understand decision-making processes. Despite these limitations, data-driven control is a prominent solution for optimizing crop yield and quality in high-tech farming systems where sensors are abundant.

In contrast, the model-based control approach in indoor farming uses mathematical models to represent the physiological and physical processes that govern plant growth and system dynamics. These models integrate knowledge about photosynthesis, respiration, transpiration, substrate water content. By simulating the crop-environment system, model-based control enables the prediction of future states and the formulation of advanced control strategies, such as Optimal Control, Model Predictive Control (MPC), etc. Model-based control has some key advantages:

- Parameters and states have physical meaning, making predictions and system behavior transparent.
- The model-based approach allow for coupling of multiple subsystems under a unified control framework.
- No need for large amounts of experimental data.

This approach also has limitations, such as complexity in development and sensitivity to uncertainties. Despite these challenges, model-based approaches remain a robust and transparent approach for controlled agriculture.

For example, [2] developed a growth model for lettuce under controlled conditions, incorporating environmental inputs such as light, temperature, and CO_2 concentration to predict dry matter accumulation, while [3] developed a stochastic MPC with chance constraints to manage greenhouse climate under uncertainty. Another lettuce growth model for control design purpose has been presented in [4].

Very recently, an accurate lettuce growth model has been developed by [5], mainly applicable to low-tech greenhouses, often affected by extremely adverse climate conditions. In particular, the model is able to account for very high and very low temperatures and humidity, allowed by low-tech greenhouse climate control to save energy and lower operating costs, while assuming that water and nutrients are adequately provided. Two parameters of the model, namely the maximum relative growth rate of dry matter at 20°, and the temperature to achieve the saturation relative growth rate, were determined by global calibration, and model performances were quantitatively evaluated by comparing the measured and simulated crop dry weights, using the root mean square error (RMSE) and relative root mean square error (RRMSE).

It must be pointed out that the [5] model focuses exclusively on greenhouse climate variables and their effects on lettuce growth, without modeling the substrate water dynamics or root-zone conditions. This is consistent with the model's purpose for optimal climate control rather than irrigation management, however, the framework is completely different for hydroponic systems, which are the focus of this work.

Unlike transpiration-based crop models such as AquaCrop [6], where water is the primary limiting factor and environmental inputs influence biomass production indirectly through reference evapotranspiration and efficiency multipliers, our photosynthesis-based model [2] treats CO_2 and light radiation as independent limiting factors that directly determine photosynthetic potential. Temperature drives distinct physiological processes–photosynthesis, growth, and respiration–rather than solely influencing evapotranspiration. This approach is particularly advantageous for controlled-environment agriculture, where environmental parameters are manageable inputs. Moreover, many factors critical to field-grown crops in AquaCrop, such as developmental stage effects, extreme temperature thresholds, and complex soil water dynamics (e.g., capillary rise) are irrelevant for indoor lettuce cultivation. While both models incorporate water balance principles, our simplified version focuses on the thin substrate layer in direct contact with the root system, reflecting the specific conditions of soilless indoor production systems.

Hydroponic systems are widely adopted in controlled environment agriculture because of their efficient use of water and nutrients and their compatibility with soilless cultivation. Hydroponic systems typically rely on inert substrates, or no substrate at all, while nutrients are supplied through a water-based solution. Therefore, in low-technology greenhouses, as reported by [5], plant growth is primarily constrained by environmental stressors such as extreme temperature and humidity, while water and nutrient availability generally remain sufficient. In contrast, hydroponic systems maintain climatic parameters within optimal ranges; in this context, the principal limitation to growth arises from water stress, which is largely determined by the efficiency and regulation of the irrigation system.

The main hydroponic techniques include:

- Nutrient Film Technique (NFT): in this system, a thin film of nutrient-rich solution flows continuously over the roots of the plants. Growth models for NFT systems, such as the one by Van Henten, often assume steady-state conditions for water and nutrient availability.
- Deep Water Culture (DWC): in this system, plants are grown with their roots submerged in a continuously oxygenated nutrient solution. This setup guarantees constant access to water and nutrients, which simplifies modeling. However, oxygen diffusion has to be accurately modeled.
- Ebb and Flow (E&F): the root zone is periodically flooded with nutrient solution and then drained. This results in a fluctuations in substrate water content, requiring models to capture this behavior.

Each technique imposes different physiological conditions on the plant, in particular regarding water availability, which must be reflected in any crop or substrate model. Despite the differences among hydroponic systems, most of the crop models in the current literature focus on general environmental variables, such as light, temperature, and CO_2 concentration, due to their high influence on crop yield. For example, [7] examine the effect of light on lettuce, while [8] focus their studies on the nutrient solution. However, water availability dynamics are often not modeled and a no water stress condition is assumed. This is acceptable in systems like DWC or NFT, where water is not a limiting factor, but problematic in systems like E&F, where water availability fluctuates over time. In these cases, modeling substrate water content and its impact on transpiration and photosynthesis is essential for predicting plant responses and managing irrigation events.

This paper focuses on the extension of the dynamic growth model proposed by [2] for lettuce (*lactuca sativa*), which is among the most commonly cultivated crops in indoor farms due to its short growth cycle. The model describes the dynamics of dry matter accumulation in lettuce based on environmental inputs such as light, temperature and CO_2 concentration. However, it does not account for the dynamics of water availability in the substrate, a factor that is important for optimizing irrigation strategies and selecting substrates. In fact, when substrate water content decreases, plants reduce water usage as a survival mechanism, which in turn results in limited photosynthesis and, consequently, limited biomass accumulation. An extension of the Van Henten model is also proposed in [9] with a main focus on the energy management of the plant, but again the paper does not address the dynamics of water content in the substrate.

The main contribution of this study is therefore the integration of substrate water content as a dynamic state variable within the existing model, enabling the simulation of its effects on plant growth. To simulate water stress effects, the dynamics of substrate water content were modeled following the procedure described by the Food and Agriculture Organization of the United Nations [10], and linked to the original model through a water stress coefficient dependent on these dynamics.

Validation of the extended model was conducted using data obtained through collaboration with Agricola Moderna, a Milan-based enterprise specializing in controlled-environment agriculture of lettuce and basil. Multiple experiments were conducted in December 2024 within the Agricola Moderna facility in Melzo, where lettuce was grown in two different substrates, peat and wood fiber, and under two different irrigation regimes. This setup was designed to induce four distinct responses of lettuce dry weight to water stress, and to create a dataset for the analyses.

Despite the water-stressed conditions, measured production was higher than that predicted by the model based on nominal parameter values, necessitating calibration. Therefore, an in-depth sensitivity and collinearity analysis was performed to identify the parameters with the greatest influence on the model. As expected, the parameter with the greatest influence was light use efficiency. This result can be explained by recalling that the Van Henten model was originally developed considering solar radiation, while in indoor agriculture light is emitted by LEDs, which do not reproduce the full solar spectrum, but mainly emit in red and blue wavelengths to optimize the growth of plants [7]. Furthermore, the light use efficiency value of the original model was based on lettuce cultivations from 1994, thus not accounting for genetic improvements introduced over the past 30 years to develop more efficient lettuce varieties. Substrate water content parameter at wilting point was also considered for calibration. After calibrating these two parameters, the model accurately reproduced lettuce dry weight production across all substrates and irrigation regimes considered in the experiments. In this respect, the Theil's inequality coefficient (TIC) [11] and the modified mean absolute relative error (MARE) coefficient, were used to assess the goodness of fit.

The new model aims to be a basis for future development of more sustainable indoor farming practices, by providing a modeling framework to guide substrate substitution and water management decisions, for example through optimal control strategies.

The paper is organized as follows. Section 2 recalls the Van Henten model. Section 3 extends the model by describing water stress conditions. Section 4 describes the implementation of the experiments. Section 5 describes the sensitivity and collinearity analysis. Section 6 reports the results of the model calibration. Section 7 draws some conclusion.

2. Van Henten model

The mathematical model developed by [2] is one of the earliest dynamic models for crop growth, and was specifically designed to describe lettuce growth in a greenhouse environment, adopting a Nutrient Film Technique (NFT) system. NFT is a hydroponic technique in which a thin film of water containing dissolved nutrients flows continuously, within a slightly sloped channel, over the bare roots of plants. In this way, water stress is avoided, due to the constant flow of nutrient solution that kept the roots well-watered.

The model, detailed in the following, simulates dry mass accumulation by modeling physiological processes such as photosynthesis, respiration, and biomass accumulation.

2.1. Input variables

The inputs of the model are the following:

- u_{PAR} [W·m^{-2}]: incident photosynthetically active radiation.
- u_T [°C]: canopy temperature.
- u_{CO_2} [ppm]: CO_2 concentration in the greenhouse air.

Photosynthetically active radiation is the portion of the light spectrum that plants can use for photosynthesis and ranges from 400 nm (blue light) to 700 nm (red light).

Canopy temperature sensors, such as infrared thermometers, can be expensive and may disturb the microclimate or the plants if placed too close. For this reason, air temperature was used as a proxy for canopy temperature. Although the two may differ throughout the day due to radiation and transpiration effects, their average values are generally close.

2.2. State variables

The model introduces two state variables[1]:

- x_{nsdw} [g·m^{-2}]: non-structural dry weight, i.e. fraction of the dry weight consisting of compounds such as glucose, sucrose, and starch.
- x_{sdw} [g·m^{-2}]: structural dry weight, accounting for the biomass of structural materials such as cell walls and cytoplasm.

The two state variables are characterized by different dynamics: a faster dynamics for x_{nsdw}, accounting for the rapid accumulation of assimilates, and a slower one for x_{sdw}, accounting for the formation of structural biomass.

2.3. Output variables

The outputs of the model are defined as:

- y_{dwp} [g]: total dry weight of the plant.
- y_{lap} [m^2]: leaf area of the plant.

Variable y_{dwp} is a key indicator of crop productivity and economic performance, as fresh weight, which is typically used for commercial evaluation, can be derived from it.

2.4. Parameters

The model is characterized by many parameters, defined below and linked to the phenomena they influence. The numerical values assumed in [2] are reported in Table A.1.

2.4.1. Photosynthesis and carbon assimilation parameters

Parameters related to CO_2 uptake, conversion and associated efficiencies.

- C_α [–]: factor that converts assimilated CO_2 into sugar (CH_2O), given by the ratio of the molecular weights of CH_2O and CO_2.
- C_ω [g·m^{-3}]: density of CO_2 (at a temperature of 15 °C and a pressure of 101.3 kPa).
- C_Γ [ppm]: CO_2 compensation point at 20 °C, i.e. the CO_2 concentration at which the rate of CO_2 uptake by photosynthesis equals the rate of CO_2 release from respiration.
- $C_{Q10,\Gamma}$ [–]: Q_{10} factor for the CO_2 compensation point, it is a coefficient that describes how the CO_2 compensation point changes with a 10 °C increase in temperature.
- C_ε [g·J^{-1}]: light use efficiency at very high CO_2 concentrations, expressed as grams of sugar produced per joule of absorbed PAR. It reflects the biochemical efficiency of photosynthesis.
- G_{bnd} [m·s^{-1}]: boundary layer conductance of lettuce leaves, representing the rate of CO_2 diffusion between the leaf surface and the surrounding air. It is assumed constant.
- G_{stm} [m·s^{-1}]: stomatal conductance, representing the rate of CO_2 exchange through the stomata. It is assumed constant.

Regarding the last two parameters, it must be recalled that in plant physiology the gas transport through the plant can be modeled through an equivalent electric circuit. In this framework, gas flux is treated as an electric current, while the voltage corresponds to the CO_2 concentration gradient, and the plant can be represented as a series of resistors that model the barriers to gas diffusion, such as the boundary layer, the stomata and the carboxylation process, i.e., the process that fixate CO_2 in the plant. The resistors vary depending on climatic conditions

[1] In Van Henten's paper the state variables are initialized with the following values: $x_{nsdw}(0) = 1.8 \times 10^{-4}$ g·m^{-2}, $x_{sdw}(0) = 5.4 \times 10^{-4}$ g·m^{-2}

and the physiological status of the plant. As a consequence, the canopy conductance to CO_2, G_{CO_2} [m·s^{-1}], can be computed as:

$$\frac{1}{G_{CO_2}} = \frac{1}{G_{car}} + \frac{1}{G_{bnd}} + \frac{1}{G_{stm}} \tag{1}$$

where: G_{car} [m·s^{-1}] is the carboxylation conductance, depending on the temperature.

Complex biochemical models exist which describe these temperature effects, but for modeling and control purposes it is assumed that the effect of temperature on the carboxylation conductance may be simplified by the following polynomial:

$$G_{car} = C_{car,1} u_T^2 + C_{car,2} u_T + C_{car,3} \tag{2}$$

2.4.2. Light interception and leaf area parameters

Parameters influencing how light is absorbed by the canopy and leaf characteristics.

- C_K [–]: extinction coefficient, it quantifies the rate at which incoming light (like PAR) is absorbed or scattered by leaves in a canopy as it penetrates downward.
- C_{lar} [m^2·g^{-1}]: structural leaf area ratio, it defines the area of leaf produced per gram of structural dry mass and is assumed to be constant.
- N_{plants} [m^{-2}]: plant density, expressed as the number of plants per square meter.

2.4.3. Growth and biomass allocation

Parameters related to biomass synthesis and partitioning.

- C_β [–]: yield factor that indicates the respiratory and synthesis losses of non-structural material due to growth.
- $C_{gr,max}$ [s^{-1}]: saturation growth rate at 20 °C, it is the maximum rate at which a plant can grow when all limiting factors (light, nutrients, water, CO_2) are supplied at non-limiting (saturating) levels.
- $C_{Q10,gr}$ [–]: Q_{10} factor for growth, it is a coefficient that describes how growth accelerates with a 10 °C increase in temperature.
- C_τ [–]: root dry weight to the total crop dry weight ratio, assumed constant.

2.4.4. Maintenance respiration

Parameters accounting for the energy cost of maintaining tissue.

- $C_{resp,sht}$ and $C_{resp,rt}$ [s^{-1}]: maintenance respiration coefficients for the shoot and root at 25 °C, expressed in the mass of glucose consumed per unit amount of structural dry matter present.
- $C_{Q10,resp}$ [–]: Q_{10} factor for respiration, it is a coefficient that describes how respiration rate changes with a 10 °C increase in temperature.

2.5. State and ouput equations

The dynamics of the non-structural and structural dry weight are defined as follows:

$$\frac{\mathrm{d}x_{nsdw}}{\mathrm{d}t} = C_\alpha f_{phot} - r_{gr} x_{sdw} - f_{resp} - \frac{1 - C_\beta}{C_\beta} r_{gr} x_{sdw} \tag{3}$$

$$\frac{\mathrm{d}x_{sdw}}{\mathrm{d}t} = r_{gr} x_{sdw} \tag{4}$$

where: f_{phot} is the gross canopy photosynthesis rate, f_{resp} is the maintenance respiration rate (both in [kg·m^{-2}·s^{-1}]) and r_{gr} [s^{-1}] is the specific growth rate.

The gross photosynthesis is defined as:

$$f_{phot} = \left(1 - \exp(-C_K C_{lar}(1 - C_\tau) x_{sdw})\right) f_{phot,max} \tag{5}$$

where: $f_{phot,max}$ [g·m^{-2}·s^{-1}] is the gross carbon dioxide assimilation rate of the canopy having an effective canopy surface of 1 m^2 per square-meter soil at complete soil covering. The term multiplying the maximum assimilation rate $f_{phot,max}$ accounts for geometrical and optical properties of the canopy with respect to incident radiation when the soil is not completely covered.

The maximum assimilation rate is computed as:

$$f_{phot,max} = \frac{\varepsilon\, u_{PAR}\, G_{CO_2}\, C_\omega\, (u_{CO_2} - \Gamma)}{\varepsilon\, u_{PAR} + G_{CO_2}\, C_\omega (u_{CO_2} - \Gamma)} \tag{6}$$

where: ε [g·J^{-1}] is the light use efficiency and Γ [ppm] is the CO_2 compensation point which accounts for photo-respiration at high light levels:

$$\Gamma = C_\Gamma \left(C_{Q10,\Gamma}\right)^{\frac{u_T - 20}{10}} \tag{7}$$

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

The specific growth rate is:

$$r_{gr} = C_{gr,max} \frac{x_{nsdw}}{x_{nsdw} + x_{sdw}} \left(C_{Q10,gr}\right)^{\frac{u_T - 20}{10}} \tag{9}$$

and finally, the maintenance respiration is:

$$f_{resp} = \left(C_{resp,sht}\,(1 - C_\tau)\, x_{sdw} + C_{resp,rt}\, C_\tau\, x_{sdw}\right)\left(C_{Q10,resp}\right)^{\frac{u_T - 25}{10}} \tag{10}$$

The two outputs of the model y_{dwp} and y_{lap} are computed as:

$$y_{dwp} = \frac{x_{nsdw} + x_{sdw}}{N_{plants}} \tag{11}$$

$$y_{lap} = \frac{(1 - C_\tau)\, C_{lar}\, x_{sdw}}{N_{plants}} \tag{12}$$

The parameters of the model are summarized in Table A.1.

3. Accounting for water stress dynamics

The model proposed by Van Henten was originally developed for controlled environments (NFT) under the assumption that water is always available in sufficient quantities such that growth is not limited. Consequently, there are no equations that describes water stress, nor how fluctuations in water availability affect plant physiology or biomass accumulation. This assumption significantly limits the applicability of the model to systems where water availability is variable.

For example, in E&F systems, like the one considered in this work, plants are placed in a substrate which is periodically irrigated by flooding and then allowed to drain. Water infiltrates the substrate by capillarity and raises the substrate water content to saturation. This results in cyclical fluctuations in water content and aeration within the substrate, that expose the roots to variable conditions of water availability and oxygen diffusion. As a consequence, plants are subjected to water stress, depending on irrigation frequency, substrate properties, and environmental demands.

Given the importance of water management in indoor hydroponics systems, particularly under reduced substrate volume and high plant density, it is necessary to extend the model. By introducing a water stress coefficient that modulates photosynthesis and growth based on substrate properties and moisture availability, the extended model can more accurately reflect the physiological limitations experienced by the plant. This adaptation improves the relevance of the model for resource-efficient production and dynamic control.

The key point is that when water becomes scarce, plants reduce water uptake through adaptive mechanisms in order to survive. The decrease in water uptake affects all previously described processes, resulting in a reduced growth rate. Therefore, it is necessary to model the dynamics of substrate water content to predict plant responses to water stress and, ultimately, crop yield:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = w_{in} - w_{out} \tag{13}$$

where: h [L] is the substrate water content of a single pot, w_{in} [L·s^{-1}] is the water inflow and w_{out} [L·s^{-1}] is the water outflow. Water inflow w_{in}

is due to irrigation and is an input of the model ($w_{in} = u_{IRR}$), while water outflow w_{out} is composed of four components: *transpiration*, *evaporation*, *percolation* and *drainage*.

Transpiration consists in the vaporization of liquid water contained in plant tissues and its subsequent release into the atmosphere and acts as a mechanism to cool the canopy when high temperatures might cause damage. Almost all the water absorbed via the roots is lost through transpiration, while only a minor fraction is used within the plant [10]. Evaporation is the process by which liquid water is converted to water vapor and removed from the evaporating surface. Direct solar radiation and, to a lesser extent, air temperature provide the energy required to the phase change from liquid to vapor. Evaporation and transpiration occur simultaneously and it is difficult to distinguish between the two processes, for this reason, these terms are combined into a single *evapotranspiration* outflow w_{et}.

Percolation refers to the downward movement of water through the substrate beyond the root zone, where the plant can no longer access it, w_p will denote the relevant outflow.

Finally, the outflow w_d accounts for the drainage when the substrate water content exceeds the field capacity.

To summarize, the water ouflow is given by:

$$w_{out} = w_{et} + w_p + w_d \tag{14}$$

each component will be defined in the following.

3.1. Evapotranspiration

The outflow rate due to evapotranspiration w_{et} can be computed as:

$$w_{et} = A \cdot \mathrm{ET}_c \tag{15}$$

where A [m^2] is the top area of a single pot ($A = a \times b$) and ET_c [mm·s^{-1}] defines the evapotranspiration[2] which, in turn, can be estimated by multiplying the reference crop evapotranspiration, ET_0 (mm·s^{-1}), by a crop coefficient, K_c [–]:

$$\mathrm{ET}_c = K_c \cdot \mathrm{ET}_0 \tag{16}$$

3.1.1. Reference crop evapotranspiration

The reference crop evapotranspiration, ET_0, is defined as the evapotranspiration of a hypothetical grass reference crop with an assumed height of 0.12 m, a fixed surface resistance of 70 s·m^{-1}, and an albedo of 0.23 (also called canopy reflection coefficient), which is a measure of how reflective a surface is. The FAO Penman-Monteith method is recommended as the sole standard for defining and computing ET_0 [10]:

$$\mathrm{ET}_0 = \frac{0.408\,\Delta\,(R_n - G) + \gamma\,\frac{900}{T+273}\,u_2 (e_s - e_a)}{\Delta + \gamma\,(1 + 0.34\,u_2)} \tag{17}$$

where Δ [kPa·°C^{-1}] is the slope of the relationship of saturation vapor pressure to air temperature, R_n and G [MJ·m^{-2}·s^{-1}] are the net radiation at the crop surface and the substrate heat flux at the substrate surface, respectively, $\gamma = 0.0672$ [kPa·°C^{-1}] is the psychrometric constant[3], a parameter that relates the energy required for water vaporization to the capacity of air to hold that vapor, u_2 [m·s^{-1}] is the mean daily wind velocity measured at 2 m above the ground, $(e_s - e_a)$ [kPa] corresponds to the vapor pressure deficit (VPD) of the air, namely the difference between the maximum vapor pressure the air can hold at a given temperature (mean saturation vapor pressure e_s) and the vapor pressure based on current humidity (actual vapor pressure e_a), T [°C] is the mean daily air temperature.

It must be pointed out that eq. (17) has been derived for open field applications, where irrigation scheduling is less frequent than daily, so that ET_0 in (17) is computed in [mm·day^{-1}]. However, since our model adopts seconds as time unit, we need to compute ET_0 in [mm·s^{-1}] [12] as well as to account for day/night transitions:

$$\mathrm{ET}_0 = \frac{0.408\,\Delta\,(R_n - G) + \gamma\,\frac{37.5}{3600\,(u_T+273)}\,u_2\,(e_s - e_a)}{\Delta + \gamma(1 + C_d\,u_2)} \tag{18}$$

where R_n and G, now in [MJ·m^{-2}·s^{-1}], and u_2 represents the instantaneous values rather than daily means, while the air temperature u_T is used instead of the daily mean temperature, C_d [s·m^{-1}] is a day/night factor which is equal to 0.24 for hours of the day and 0.96 for hours of the night. To model the transition from day (u_{PAR} high) to night (u_{PAR} low) a logistic function is introduced:

$$f_{\mathrm{day/night}}(u_{PAR}) = \frac{1}{1 + \exp(-K\,(\eta - u_{PAR}))} \tag{19}$$

where η is the PAR threshold below which the night condition is assumed ($\eta = 10\,\mathrm{W \cdot m^{-2}}$) and K is a parameter that regulates the steepness of the transition ($K = 200\,\mathrm{W^{-1} \cdot m^2}$), therefore, C_d is computed as follows:

$$C_d = 0.24 + 0.72\,f_{\mathrm{day/night}}(u_{PAR}) \tag{20}$$

The net radiation at the crop surface R_n is given by the difference between the incoming net shortwave radiation R_{ns} and the outgoing net longwave radiation R_{nl}, negligible in indoor environments with respect to R_{ns}, the latter given by:

$$R_{ns} = (1 - \alpha) R_s \tag{21}$$

in open field environments, where α is the albedo and R_s is the incoming solar radiation. In indoor farming, R_s is replaced by the radiation emitted by the LED lights, i.e. the model input u_{PAR}. Furthermore, the walls of the indoor farming environment reflect the light reflected by the plant, resulting in $\alpha \approx 0$, so that we can write:

$$R_n = R_{ns} = 10^{-6}\,u_{PAR} \tag{22}$$

where the factor 10^{-6} is needed to convert u_{PAR} [W·m^{-1}] in [MJ·m^{-2}·s^{-1}]

The substrate heat flux at the substrate surface is approximated as a fraction of R_n:

$$G = C_G\,R_n = 10^{-6}\,C_G\,u_{PAR} \tag{23}$$

where C_G is equal to 0.1 during daylight periods and 0.5 during nighttime periods [10], hence:

$$C_G = 0.1 + 0.4\,f_{\mathrm{day/night}}(u_{PAR}) \tag{24}$$

The mean saturation vapor pressure, e_s, is computed as the average of the saturation vapor pressure at the maximum air temperature, $u_{T,\max}$, and the minimum air temperature $u_{T,\min}$ during the period from transplant to harvest, estimated as $u_{T,\max} = 23\,°\mathrm{C}$ and $u_{T,\min} = 20\,°\mathrm{C}$:

$$e_s = \frac{e^o(u_{T,\max}) + e^o(u_{T,\min})}{2} \tag{25}$$

where e^o (kPa) is the saturation vapor pressure, namely the pressure at which water vapor is in thermodynamic equilibrium with its condensed state [10]:

$$e^o(u_T) = 0.6108\,\exp\left(\frac{17.27\,u_T}{u_T + 237.3}\right) \tag{26}$$

The actual vapor pressure e_a is the partial pressure of water vapor in the air, representing the amount of water vapor in a given volume and it is computed from $e^o(T)$ and the relative humidity u_{RH}, which is a model input:

$$e_a(u_T) = \frac{e^o(u_T)\,u_{RH}}{100} \tag{27}$$

The slope of the relationship of saturation vapor pressure to air temperature Δ is given by:

$$\Delta(u_T) = \frac{4098\,e^o(u_T)}{(u_T + 237.3)^2} \tag{28}$$

Finally, u_2 is also defined as a model input.

[2] 1 [mm] corresponds to 1 [L·m^{-2}].

[3] Value of the psychrometric constant at atmospheric pressure at sea level (101.3 kPa).

3.1.2. Crop coefficient

Two approaches can be adopted to model the crop coefficient: single and dual crop coefficient [10]. In the single crop coefficient approach, the effects of crop transpiration and substrate evaporation are combined into a single K_c coefficient, while in the dual crop coefficient approach, the effects of crop transpiration and substrate evaporation are determined separately using the relationship:

$$K_c = K_e + K_{cb} \tag{29}$$

where K_{cb} is the basal crop coefficient and K_e is the substrate water evaporation coefficient. The basal crop coefficient represents transpiration under optimal conditions and it is defined as the ratio of the crop evapotranspiration to the reference evapotranspiration (e_c/e_0) when the substrate surface is dry but transpiration is occurring without any limitation due to water availability. The substrate water evaporation coefficient accounts for substrate surface evaporation. In this work the dual crop coefficient approach is adopted, since it gives more accurate estimates for K_c.

The basal crop coefficient is computed using the formula proposed by FAO [10]:

$$K_{cb} = K_{cb,TAB} + [0.04\,(u_2 - 2) - 0.004\,(u_{RH,\min} - 45)]\left(\frac{h_{c,\max}}{3}\right)^{0.3} \tag{30}$$

where $h_{c,\max}$ is the maximum crop height for the crop, namely 0.3 m for lettuce according to FAO, $K_{cb,TAB}$ is the value of K_{cb} for non stressed, well-managed crops in sub-humid climates, namely 0.9 for lettuce according to FAO, $u_{RH,\min}$ is the minimum relative humidity during the considered time period. Since relative humidity changes only from day to night, while keeping constant values during these periods, u_{RH} is assumed instead of $u_{RH,\min}$.

Evaporation is the process by which water vapor escapes from substrate pores into the atmosphere, and occurs predominantly from the exposed substrate fraction. Hence, evaporation is restricted at any moment by the energy available at the exposed substrate fraction. Following a wetting event, K_e reaches its maximum value and can be defined as follows [10]:

$$K_e = \min(K_r\,(K_{c,\max} - K_{cb}),\, f_{ew}\,K_{c,\max}) \tag{31}$$

where K_r is a dimensionless evaporation reduction coefficient, K_{cb} is the basal crop coefficient, $K_{c,\max}$ is the maximum value of K_c following rain or irrigation and f_{ew} is the fraction of substrate that is both exposed and wetted.

In general, the evaporation process consists of two stages: an initial phase with a relatively high and constant rate (stage 1), followed by a second phase where the evaporation rate gradually decreases (stage 2). To model this behavior the coefficient K_r can be computed as follows:

$$K_r = \begin{cases} 1 & \text{if } D_e \le REW \\ \frac{TEW - D_e}{TEW - REW} & \text{if } REW < D_e \le TEW \\ 0 & \text{if } D_e > TEW \end{cases} \tag{32}$$

where TEW [mm] is the Total Evaporable Water, i.e. the maximum depth of water that can be evaporated from the topsoil after full wetting, REW [mm] is the Readily Evaporable Water, i.e., the cumulative depth of evaporation at the end of stage 1, D_e [mm] is the cumulative depth of evaporation (depletion) from the topsoil. When D_e is smaller than REW evaporation remains in stage 1 and K_r keeps its maximum value of 1, as D_e increases beyond REW, K_r decreases linearly until it reaches 0 at $D_e = TEW$, after this point evaporation stops.

The value of TEW can be computed as:

$$TEW = 1000\,(\theta_{FC} - 0.5\,\theta_{WP})z_e \tag{33}$$

where θ_{FC} [–] is the volumetric water content at field capacity, θ_{WP} [–] is the volumetric water content at wilting point, z_e [m] is the depth of the topsoil. Since in E&F the substrate's depth does not exceed 10 cm, we made the assumption that the whole substrate is topsoil, so $z_e = z$. For peat and wood fiber, the values of θ_{FC} and θ_{WP} were computed experimentally as reported in Section 4. The value of REW is computed as [13]:

$$REW = \min\left(2 + \frac{TEW}{3}, 0.8\,TEW\right) \tag{34}$$

The cumulative depletion D_e can be approximated as the difference between θ_{FC} and the substrate water content in volumetric units θ

$$\theta = \frac{h}{1000\,V} \tag{35}$$

multiplied by the depth of the substrate z:

$$D_e = (\theta_{FC} - \theta)\,z \tag{36}$$

where V [m^3] is the volume of a single pot (if the pot is a prism $V = a \times b \times z$[4]).

The coefficient $K_{c,\max}$ is computed the following way [10]:

$$K_{c,\max} = \max\left(1.2 + K_{cb} - K_{cb,TAB}, K_{cb} + 0.05\right) \tag{37}$$

while the fraction of substrate that is both exposed and wetted f_{ew} is computed as:

$$f_{ew} = \min(1 - f_c, f_w) \tag{38}$$

where: $1 - f_c$ is the fraction of substrate that is not shaded and f_w is the fraction of substrate wetted by irrigation. In an E&F system, $f_w = 1$ can be assumed and, consequentially:

$$f_{ew} = 1 - f_c \tag{39}$$

Finally, f_c is computed as:

$$f_c = \min(1, 1 - \exp(-C_K\,C_{lar}\,(1 - C_\tau)\,x_{sdw})) \tag{40}$$

where the exponential term is the same that accounts for geometrical and optical properties of the canopy used by [2] in his model.

3.1.3. Water stress

However, the effect of water stress has not yet been described, in fact, when the substrate water content decreases, transpiration decreases too. This behavior can be taken into account through a water stress coefficient K_s [–]. It is often assumed that K_s decreases linearly with the substrate water content, however, this linear assumption often oversimplifies the actual plant response. For this reason, the coefficient K_s is modeled as suggested by [14]:

$$K_s = \begin{cases} 0 & \text{if } \theta < \theta_{WP} \\ \left(\frac{\theta - \theta_{WP}}{\theta_{FC} - \theta_{WP}}\right)^q & \text{if } \theta_{WP} \le \theta < \theta_{FC} \\ 1 & \text{if } \theta \ge \theta_{FC} \end{cases} \tag{41}$$

where q is a parameter that determines the degree of curvature of the relation between substrate water content and transpiration. Since lettuce is very sensitive to water stress [15], we set $q = 1.5$, reflecting that even a small loss of substrate water content results in a greater reduction in transpiration compared to a linear decline. Therefore:

$$ET_c = (K_e + K_s\,K_{cb})\,ET_0 \tag{42}$$

3.2. Drainage

The drainage term w_d is zero when $\theta > \theta_{FC}$ and constant to a value δ otherwise. The parameter δ should be large enough to return θ to field capacity within a few minutes, in this work it is set to 0.06 [$L{\cdot}s^{-1}$]. A logistic function is again used to model the switch:

$$w_d = \frac{\delta}{1 + \exp(-K_d\,(\theta - \theta_{FC}))} \tag{43}$$

where K_d is a parameter that regulates the speed at which the condition switches; it is set to 20,000 [$L^{-1}{\cdot}s$] since we want drainage to stop immediately when θ reaches field capacity.

[4] Data not disclosed.

3.3. Percolation

The percolation rate (of a single pot) w_p was modeled as a constant term that is continuously subtracted from the total water content of the substrate, depending on the water retention properties of the substrate, namely:

$$w_p = A \cdot P \tag{44}$$

with P [L·m^{-2}·s^{-1}] being the percolation rate, determined experimentally for peat and wood fiber as reported in Section 4.

3.4. Growth model extension

However, to adapt Van Henten's model to indoor farming conditions, some further modifications are necessary.

3.4.1. Extinction coefficient

Van Henten set the extinction coefficient C_K to 0.9, a typical value for planophile canopies [2]. Since lettuce has nearly horizontal leaves, this value initially appeared reasonable. However, the coefficient was measured under natural sunlight, in indoor farming the fixed position of artificial light sources and the presence of reflective or absorptive materials significantly alter the light distribution. In addition, high plant density in indoor farming influences leaf orientation, causing leaves to grow more vertically compared to low-density conditions. As a result, the actual value of the extinction coefficient is lower, based on measurements reported by [16] it was set $C_K = 0.6$.

3.4.2. Boundary layer conductance

Boundary layer conductance in Van Henten's work was computed under specific climate conditions, therefore it has been reevaluated using the formulation proposed by [5] for the boundary layer resistance R_{bnd} [s·m^{-1}]:

$$R_{bnd} = (Le)^{0.67} \frac{1174\,(L_f)^{0.5}}{\left(L_f\,|T_{air} - T_{canopy}| + 207\,u_2^2\right)^{0.25}} \tag{4.35}$$

where T_{air} is the air temperature, T_{canopy} is the temperature measured at canopy level, Le [–] is the Lewis number for CO_2, defined as the ratio of thermal diffusivity to mass diffusivity, for CO_2 in air at 25 °C $Le = 1.47$, L_f [m] is the leaf characteristic dimension, taken as the mean leaf width in the wind direction. Setting again the difference between the two temperatures to zero and $L_f = 0.075$ m, as suggested by Van Henten, the boundary layer conductance is given by:

$$G_{bnd} = \frac{(207\,u_2^2)^{0.25}}{1174\,(Le)^{0.67}\,L_f^{0.5}} \tag{4.36}$$

3.4.3. Water stress and growth

Finally we need to model the influence of substrate water content on plant growth. Practically, this means introducing a water stress coefficient that limits one of the physiological processes contributing to plant growth. A brief explanation of the plant's response mechanisms under stress is needed to justify our modeling approach.

When exposed to abiotic or biotic stress, such as water deficit, plants produce abscisic acid (ABA), a hormone that induces stomatal closure. This response reduces transpirational water loss and limits CO_2 uptake, as a result, photosynthesis is inhibited. Other effects of water stress include reductions in the maximum carboxylation rate, the maximum photosynthetic electron transport rate, and mesophyll conductance [17].

Initially, only the stomatal conductance G_{stm} was multiplied by the water stress coefficient K_s. However, the results showed that the limiting effect was too small, suggesting that restricting only the stomatal conductance was insufficient. For this reason, the effect of water stress was modeled by multiplying f_{phot} by K_s, effectively limiting the entire photosynthetic process. This simplification is commonly adopted, as noted by [14] in their analysis of land surface models.

3.5. Extended model summary

Collecting all previous definitions, the growth model accounting for water stress dynamics is characterized by 6 inputs:

- u_{PAR} [W·m^{-2}]: incident photosynthetically active radiation.
- u_T [°C]: air (canopy) temperature.
- u_{CO_2} [ppm]: CO_2 concentration in the greenhouse air.
- u_{RH} [–]: relative humidity.
- u_2 [m·s^{-1}]: wind speed at 2 m height.
- u_{IRR} [L·s^{-1}]: irrigation inflow.

3 states:

- x_{nsdw} [g·m^{-2}]: non-structural dry weight.
- x_{sdw} [g·m^{-2}]: structural dry weight.
- h [L]: substrate water content of a single pot.

2 outputs:

- y_{dw} [kg·m^{-2}]: total dry weight over 1 m^2.
- y_{la} [m^2·m^{-2}]: leaf area over 1 m^2.

However, of the two outputs, only y_{dw} was measured and used for validation, while y_{la} was excluded from the model analysis. This choice is justified by the fact that dry weight is the most relevant output in terms of economic return.

The state and output equations are finally given by:

$$\frac{dx_{nsdw}}{dt} = C_\alpha\,K_s\,f_{phot} - r_{gr}\,x_{sdw} - f_{resp} - \frac{1-C_\beta}{C_\beta}\,r_{gr}\,x_{sdw} \tag{45}$$

$$\frac{dx_{sdw}}{dt} = r_{gr}\,x_{sdw} \tag{46}$$

$$\frac{dh}{dt} = u_{IRR} - w_{et} - w_p - w_d \tag{47}$$

$$y_{dw} = x_{nsdw} + x_{sdw} \tag{48}$$

$$y_{la} = (1 - C_\tau)\,C_{lar}\,x_{sdw} \tag{49}$$

The same parameters defined for the Van Henten model, reported in Table A.1, were used, except for C_K, set to 0.6, and G_{bnd}, which is now a variable that depends on u_2. The added parameters are shown in Table B.2, a block scheme of the model is shown in Fig. 1, while the variables of the model are listed in Tables C.3 and C.4 in Appendix C.

4. Experimental setup

To support the validation and calibration of the extended model, a series of experiments were carried out at the Agricola Moderna facility in Melzo (Italy) in December 2024 (Fig. 2).

Lettuce grown in peat and wood fiber substrates was subjected to two different irrigation regimes, and growth was monitored over time. In addition, some substrate properties that were not available in the literature, in particular water retention characteristics such as substrate water content at field capacity and percolation rate, were experimentally determined.

On November 20th 2024, seedlings were planted in a peat substrate and grown in a growing room for 15 days, during this initial period all plants were subjected to the same treatment and conditions.

On December 5th, a total of 420 plants were selected exhibiting the healthiest characteristic, 210 plants were placed on Bench 1, which was scheduled to flood three times per day, and the remaining plants were placed on Bench 2, which was scheduled to flood once per day (Fig. 3). In turn, each bench was divided in 2 trays of dimensions a, b, and z, representing the length, the width and the depth of a single pot respectively, one containing peat as substrate and the other wood fiber, 105 plants were placed in each tray, resulting in a total of 4 treatments. The same nutrient solution was applied across all treatments. Labels P3, P1, W3, and W1 in Fig. 3 denote the four treatments: peat with three irrigations per day, peat with one irrigation per day, wood fiber with three

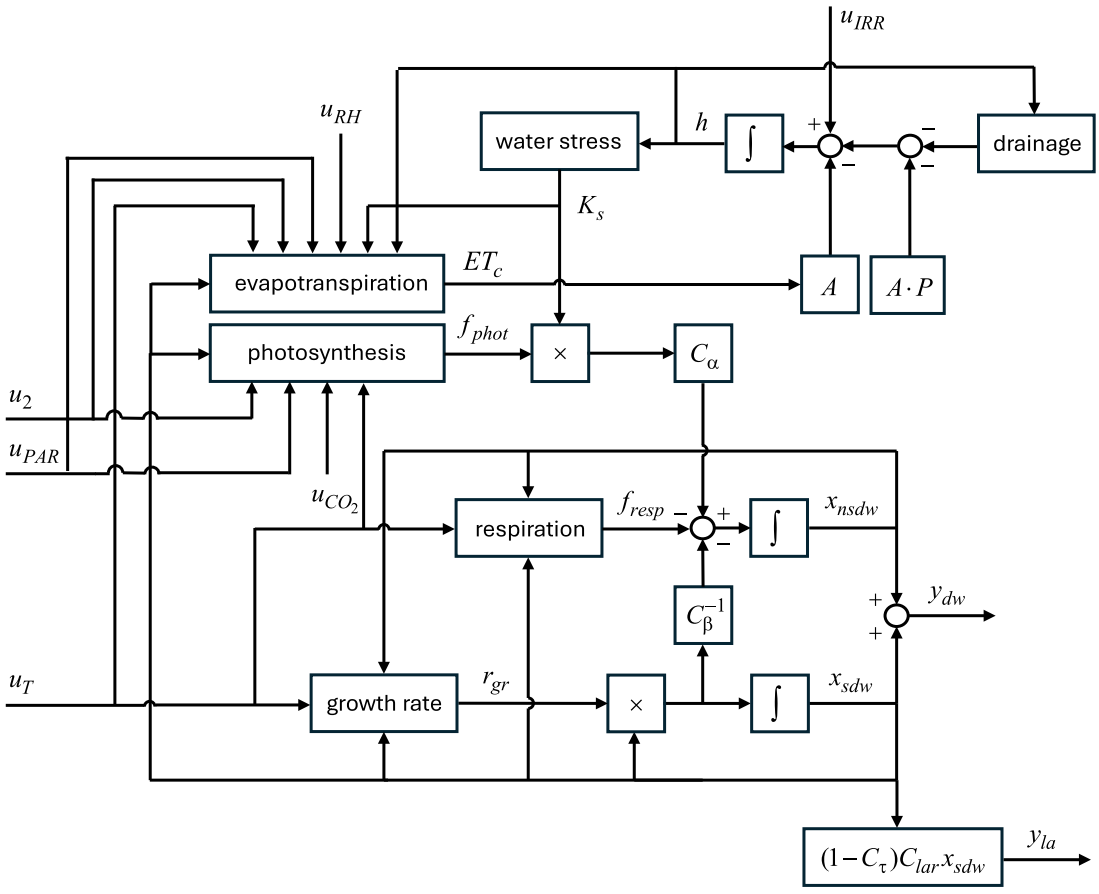

Fig. 1. Block scheme of the model.

Fig. 2. Ebb and Flow irrigation system at Agricola Moderna.

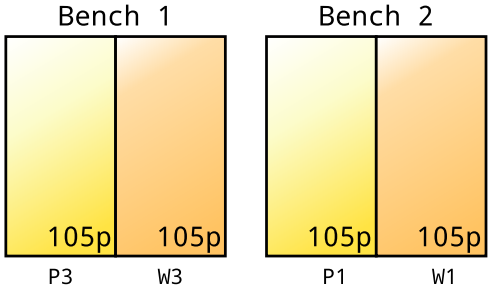

Fig. 3. Experimental benches.

irrigations per day, and wood fiber with one irrigation per day, respectively. In addition, the fresh weight of 37 plants that were not selected for transplantation were measured to determine the initial conditions for x_{nsdw} and x_{sdw}:

$$x_{nsdw}(0) = 7.3876 \times 10^{-4}\,\mathrm{Kg}\cdot\mathrm{m}^{-2} \tag{50}$$

$$x_{sdw}(0) = 2.21628 \times 10^{-3}\,\mathrm{Kg}\cdot\mathrm{m}^{-2} \tag{51}$$

$$\theta(0) = \theta_{FC} = \begin{cases} 0.342 & \text{(peat)} \\ 0.312 & \text{(wood)} \end{cases} \tag{52}$$

In particular, the initial values for x_{nsdw} and x_{sdw} were determined as the 25 % and 75 % of the mean over the 37 measurements of y_{dw} detected on the day of the transplant, while the value for θ_{FC} was determined as explained in the following.

The input u_{PAR} was provided by a LED lighting system, emitting light in the PAR portion of the spectrum, the LEDs were on during the day, producing 47.78 $\mathrm{W}\cdot\mathrm{m}^{-2}$, and off at night, producing 0 $\mathrm{W}\cdot\mathrm{m}^{-2}$ (Fig. 4).

The air temperature u_T was measured every 5 minutes by a temperature sensor positioned above the benches (Fig. 5).

The u_{CO_2} concentration was maintained constant at approximately 800 ppm throughout the experiment[5]

The relative humidity u_{RH} was regulated to 55 % during the day and 75 % during the night (Fig. 6).

The wind speed in a controlled environment can be considered nearly constant and actually in the experiment varied in the range 0.3 $\mathrm{m}\cdot\mathrm{s}^{-1}$ and 0.5 $\mathrm{m}\cdot\mathrm{s}^{-1}$, so u_2 has been set constant to a mean value of 0.4 $\mathrm{m}\cdot\mathrm{s}^{-1}$.

The irrigation flow rate was assigned a pulse pattern, with an amplitude such as to immediately restore θ to the field capacity; $u_{IRR} = 0.11$

[5] PAR and CO_2 levels were more stable in our indoor experiment compared to those measured by Van Henten during his study.

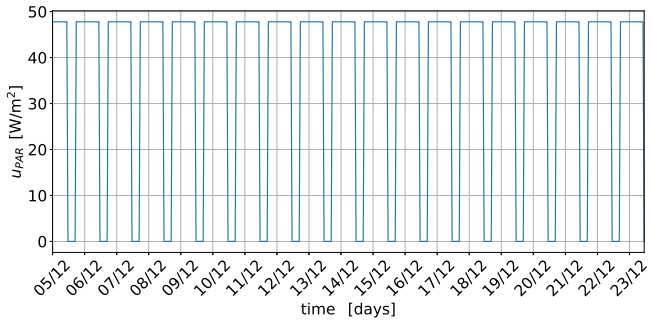

Fig. 4. Photosynthetically Active Radiation during the experiment.

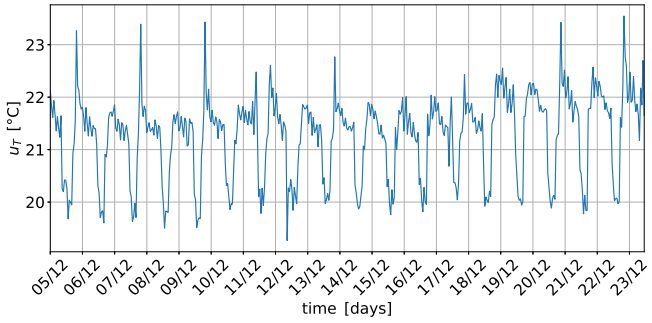

Fig. 5. Air temperature during the experiment.

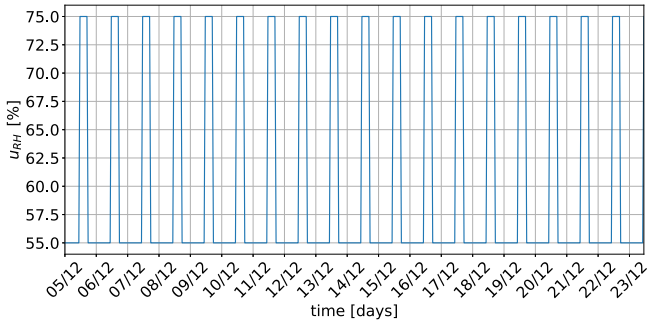

Fig. 6. Relative humidity during the experiment.

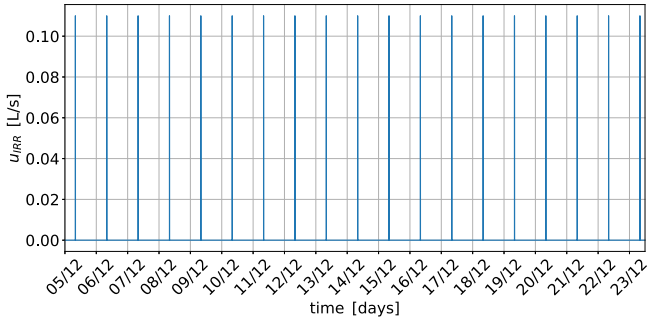

Fig. 7. Irrigation inflow during the experiment.

L/s was chosen. As already mentioned, the frequency of these pulses is equal to 1 or 3 times a day and the duration of the pulse is equal to one hundredth of the period (Fig. 7 shows u_{IRR} with a period of 1 day).

On December 12th, 16th, 18th, and 20th, 15 plants from each treatment were harvested by cutting them at the stem base (at substrate surface level). Fresh weight of each plant were recorded and stored the data in a database. On December 23rd, 30 plants were harvested and the remaining ones discarded. Then, for each sample, the average fresh weight of the harvested plants was computed, to account for root biomass this

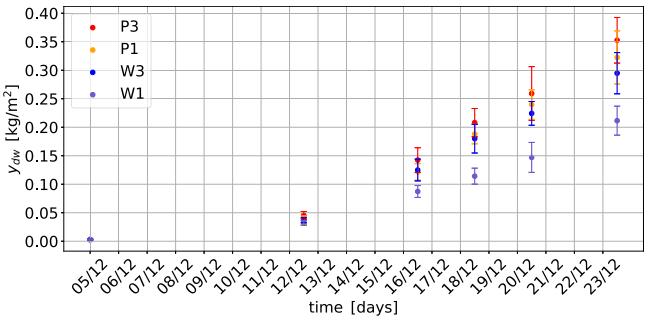

Fig. 8. Measured dry weight y_{dw} among P1, P3, W1, and W3 with error bars.

value was multiplied by 1.15, in accordance with [2]. Based on a previous experiment conducted by Agricola Moderna, which found that dry weight corresponds to 4 % of the total fresh weight, dry weight was estimated by multiplying the adjusted fresh weight by 0.04.

It must be pointed out that in the Van Henten model the output variable y_{dwp} is expressed in grams and represents the dry weight of a single plant, while in this work the dry weight per square meter y_{dw} (in kg·m^{-2}) is considered, therefore, the resulting weight was multiplied by the plant density.

Fig. 8 shows the resulting dry weight estimates for the four treatments, error bars show the standard deviation over the 15 samples (30 for the last). Among treatments, P3 achieved highest yield, furthermore, water stress had a greater effect on wood fiber substrates, as the dry weights of W3 and W1 differ by approximately 0.2 kg, while the difference is smaller for peat treatments.

Statistical analyses were performed in Python (scipy, scikit-posthocs, statsmodels) to assess the significance of the differences among the four treatments, and the significance level α was set at 0.05.

Data distributions of each group were firstly assessed using the Shapiro-Wilk test (p-values calculated separately for each treatment), and homogeneity of variances was evaluated with Levene's test [18]. These preliminary statistical checks indicated that the four treatment datasets did not meet the assumptions of normality and homoscedasticity. In fact, the Shapiro-Wilk test showed significant deviations from normality for all treatments (P1: W(Shapiro-Wilk test statistic) = 0.960, $p = 0.019$; W1: $W = 0.952$, $p = 0.006$; P3: $W = 0.950$, $p = 0.005$; W3: $W = 0.948$, $p = 0.004$), and Levene's test confirmed unequal variances among groups (Levene's statistic = 7.391, $p < 0.001$). As normality conditions were not always satisfied a non-parametric approach was considered the most appropriate.

First, the Kruskal-Wallis test revealed a significant overall difference among treatments (H(Kruskal-Wallis test statistic) = 30.611, $p = 0.000001$), then Dunn-Bonferroni post-hoc analysis highlighted a clear and consistent pattern: W1 was the only treatment significantly different from all others. Specifically, W1 differed from P1 ($p = 0.000338$), from P3 ($p = 8.3 \times 10^{-7}$), and from W3 ($p = 0.002523$). In contrast, P1, P3, and W3 did not differ significantly from each other, with all pairwise comparisons yielding non-significant p-values (all corrected p-values ≥ 0.49). Overall, this pattern indicates that the behaviour of W1 is not only numerically distinct but also statistically robust. The consistent separation of W1 from all other treatments strengthens the interpretation that the observed differences are unlikely to result from random variability, whereas the overlap among P1, P3, and W3 suggests a comparable response across these three treatments.

To model the dynamics of the substrate water content, some substrate-dependent parameters are required. One of the most important parameters is the volumetric Total Available Water (TAW), defined as the difference between the substrate water content at field capacity θ_{FC}, and the substrate water content at the wilting point θ_{WP}.

As mentioned, the substrate water content at field capacity θ_{FC} is the water content held by the substrate after excess water has drained away

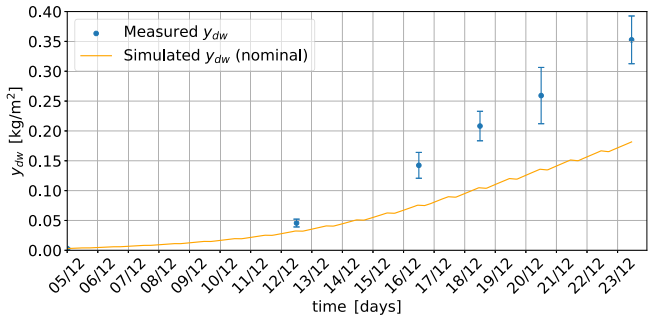

Fig. 9. P3: uncalibrated, simulated nominal dry weight and measured dry weight.

and the rate of downward movement has significantly decreased, it was measured for each substrate using a volumetric water content sensor placed in the substrate (HOBO MX230x). After the substrate reached saturation through irrigation, it was allowed to dry naturally to residual levels, then, by analyzing the slope change in the water content curve, the point corresponding to θ_{FC} was identified for both peat and wood fiber: for peat, θ_{FC} was set to 34.2 %, while for wood fiber it was set to 31.2 %.

The substrate water content at wilting point θ_{WP} corresponds to the minimum water content the plant requires not to wilt. Since we did not have access to instrumentation to measure θ_{WP}, the values reported in the substrate producers' data sheets were used to derive an initial assumption for the actual values. The reported values are: for peat, $\bar{\theta}_{FC} = 69.43\,\%$ and $\bar{\theta}_{WP} = 35.66\,\%$, for wood fiber, $\bar{\theta}_{FC} = 45.65\,\%$ and $\bar{\theta}_{WP} = 16.68\,\%$. The actual θ_{WP} for peat was assumed to maintain the same proportion with respect to θ_{FC} as indicated in the datasheet, and set to 17.57 %. From the datasheet, the ratio of TAW between the two substrates is given by:

$$\frac{\overline{TAW}_{\text{wood}}}{\overline{TAW}_{\text{peat}}} = \frac{(\bar{\theta}_{FC} - \bar{\theta}_{WP})_{\text{wood}}}{(\bar{\theta}_{FC} - \bar{\theta}_{WP})_{\text{peat}}} = \frac{0.4565 - 0.1668}{0.6943 - 0.3566} = 0.8578 \tag{53}$$

then, this ratio was applied to the measured θ_{FC} value to estimate θ_{WP} for wood fiber: using $\theta_{FC} = 34.2\,\%$ and $\theta_{WP} = 17.57\,\%$ for peat, TAW_{peat} is 16.63 %, assuming $TAW_{\text{wood}} = 85.78\,\%\,(TAW_{\text{peat}})$, $\theta_{WP} = 16.93\,\%$ was estimated. It must be pointed out that this method provides only an initial estimate of θ_{WP} for both substrates, the actual values were later determined through calibration.

Another parameter that was measured was the percolation rate P, which represents the rate at which water moves downward beyond the root zone, in indoor farming this corresponds to water leaking from the pot after field capacity is reached.

To measure P this procedure was followed: a tray was filled with the substrate, wetted, and placed inside a dark room to prevent evaporation due to radiation. The tray was weighted when the substrate was at field capacity and then weighed again 24 hours later. Since 1 kg of water corresponds to 1 L, the difference in weight corresponds to the volume of water lost through percolation. This experiment was repeated separately for peat and wood fiber substrates.

The resulting daily percolation rates were 0.421 L·m^{-2}·day^{-1} for peat and 0.875 L·m^{-2}·day^{-1} for wood fiber. Assuming that percolation occurs at a constant rate, P was computed by dividing the daily values by 86400 s·day^{-1}, obtaining 4.871×10^{-6} L·m^{-2}·s^{-1} and 1.0127×10^{-5} L·m^{-2}·s^{-1} for peat and wood fiber, respectively.

Fig. 9 shows a comparison between simulated dry weight in the nominal conditions and measured dry weight for treatment P3, similar results were obtained for the other treatments: it is evident that the model needs calibration, especially to replicate the increased productivity obtained despite water stress conditions.

5. Sensitivity and collinearity analysis

Since the extended model is characterized by 37 parameters, a sensitivity analysis is needed to select a vector of *practically identifiable parameters* [19], having the largest effect on the output of the model, while also avoiding collinearity among parameters.

A number of 20 parameters have been excluded from the analysis:

- C_α, G_{bnd} and C_ω have been already defined as constant.
- C_τ: defined as constant to ensure consistency in the estimation of total biomass across all simulations, the shoot-to-root ratio has been used to compute the total dry weight from the shoot dry weight.
- $C_{Q10,\Gamma}$, $C_{Q10,gr}$, $C_{Q10,resp}$: since temperature remained nearly constant in the experiments, these parameters have been considered as constant across all physiological models.
- a, b, z, N_{plants}: parameters defining the tray dimensions and planting density.
- Le, γ, $K_{cb,TAB}$, $h_{c,\max}$, α: values assumed from the literature.
- η, δ, K_{d}, K: these parameters do not have a physical interpretation and were introduced solely to shape the transition behavior in the model.

The remaining $m = 17$ parameters, shown in red and blue in Tables A.1 and B.2, and collected in the vector $\boldsymbol{\theta} \in \mathbb{R}^m$, have been considered for sensitivity and collinearity analysis, following the approach described in [20] and here briefly recalled. The parameters shown in blue have been considered for model calibration.

For the sake of simplicity, from now on $y(t_i)$ will denote the measurement of the dry weight y_{dw} at time t_i, and $\mathbf{y} \in \mathbb{R}^n$ the vector collecting all measurements, while $\tilde{y}(t_i, \boldsymbol{\theta})$ will denote the model output $\tilde{y}_{dw}$ at time t_i, assuming the vector of parameters $\boldsymbol{\theta}$.

To identify a set of p practically identifiable parameters two indices have been considered: the *parameter importance index* ρ_j (mean square sensitivity index), measuring the sensitivity of each parameter θ_j on $\tilde{y}_{dw}$, and the *collinearity* index γ_P.

First, we define the sensitivity matrix $\mathbf{S}(\boldsymbol{\theta}) \in \mathbb{R}^{n \times m}$, whose elements are given by:

$$s_{i,j}(\boldsymbol{\theta}) = \frac{\partial \tilde{y}(t_i, \boldsymbol{\theta})}{\partial \theta_j} \tag{54}$$

with $\tilde{y}(t_i, \boldsymbol{\theta})$ being the model output at time t_i, assuming the vector of parameters $\boldsymbol{\theta}$. In particular, the sensitivity matrix $\mathbf{S}(\bar{\boldsymbol{\theta}})$ is numerically computed through a finite difference approximation:

$$s_{i,j}(\bar{\boldsymbol{\theta}}) = \frac{\partial \tilde{y}(t_i, \bar{\theta}_j)}{\partial \theta_j} \approx \frac{\tilde{y}(t_i, \bar{\theta}_j + \Delta\bar{\theta}_j) - \tilde{y}(t_i, \bar{\theta}_j - \Delta\bar{\theta}_j)}{2\,\Delta\bar{\theta}_j} \tag{55}$$

where $\bar{\boldsymbol{\theta}}$ is the vector of the nominal values of the parameters (Tables A.1 and B.2) and $\Delta\bar{\theta}_j$ is a 5 % perturbation of the parameter $\bar{\theta}_j$. In turn, the scaled sensitivity matrix $\hat{\mathbf{S}}(\bar{\boldsymbol{\theta}}) \in \mathbb{R}^{n \times m}$ is defined, whose elements are given by:

$$\hat{s}_{i,j}(\bar{\boldsymbol{\theta}}) = \frac{\bar{\theta}_j}{\bar{\mathbf{y}}}\, s_{i,j}(\bar{\boldsymbol{\theta}}) \tag{56}$$

where $\bar{\mathbf{y}}$ is the mean value of measurements:

$$\bar{\mathbf{y}} = \frac{1}{n} \sum_{i=1}^{n} y(t_i) \tag{57}$$

The parameter importance indices have been computed as:

$$\rho_j(\bar{\boldsymbol{\theta}}) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \hat{s}_{i,j}^2(\bar{\boldsymbol{\theta}})} \tag{58}$$

as an example, the ranking for treatment P3 is shown in Fig. 10.

For all four treatments C_ε was the top ranked parameter ($\rho_{C_\varepsilon} = 1.47, 1.32, 1.51$ for P1, W3 and W1 respectively), while slight variations occur in the position of the other parameters.

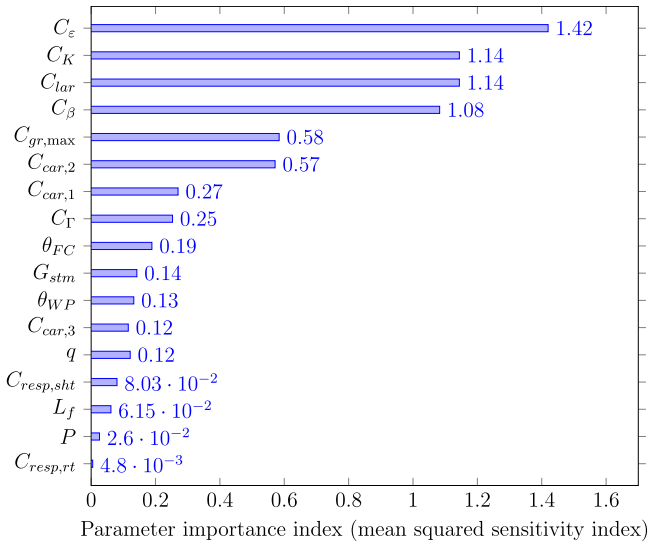

Fig. 10. P3: ranking of parameter importance.

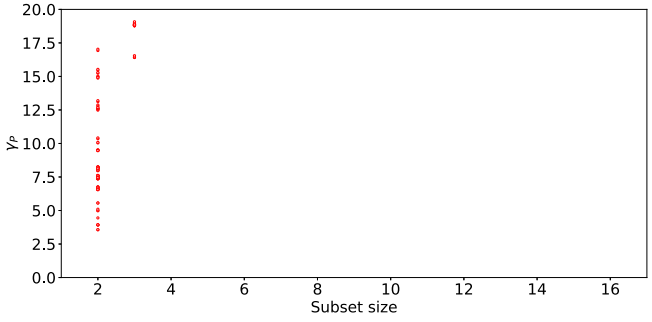

Fig. 11. P3: collinearity indices.

In this respect, it must be emphasized that the nominal value of C_ϵ from the Van Henten model was originally computed considering solar radiation. However, in indoor agriculture, light is emitted by LEDs, which do not reproduce the full solar spectrum, but mainly emit in red and blue wavelengths to optimize the growth of plants, as reported by [7]. Furthermore, the nominal value was based on lettuce cultivations from 1994; over the past 30 years, genetic improvements have been made to develop more efficient lettuce varieties. For this reason, we expected the actual value of C_ϵ to be higher and decided to recalibrate it, allowing the calibration process to capture all such changes.

In oder to select other parameters to calibrate, a collinearity analysis was performed, by computing the collinearity index $\gamma_P(\bar{\boldsymbol{\theta}})$ for each subset P of p parameters of the whole parameter vector $\boldsymbol{\theta}$:

$$\gamma_P(\bar{\boldsymbol{\theta}}) = \frac{1}{\sqrt{\min\left[\mathrm{eig}\left(\mathbf{A}_P^T(\bar{\boldsymbol{\theta}})\,\mathbf{A}_P(\bar{\boldsymbol{\theta}})\right)\right]}} \tag{59}$$

where $\mathbf{A}_P(\bar{\boldsymbol{\theta}})$ is the normalized scaled sensitivity matrix computed for the generic subset P of p parameters, whose columns $\mathbf{a}_{P,k}(\bar{\boldsymbol{\theta}})$ are given by:

$$\mathbf{a}_{P,k}(\bar{\boldsymbol{\theta}}) = \frac{\hat{\mathbf{s}}_j(\bar{\boldsymbol{\theta}})}{\|\hat{\mathbf{s}}_j(\bar{\boldsymbol{\theta}})\|} \quad \begin{cases} k = 1, \dots, p \\ j \in P \end{cases} \tag{60}$$

where $\hat{\mathbf{s}}_j(\bar{\boldsymbol{\theta}})$ is the column of the scaled sensitivity matrix $\hat{\mathbf{S}}(\bar{\boldsymbol{\theta}})$. Fig. 11 shows a scatter plot of collinearity indices for treatment P3, similar results were obtained for the other treatments.

The ordinate axis has been limited to 20, since the subset of practical identifiable parameters should be characterized by a value of γ_P lower than a critical value of 20, according to [19]. Moreover, from Fig. 11 it is evident that the parameters of the extended model exhibit high collinearity overall, and only parameter pairs can be considered practically identifiable.

The choice on the second parameter to consider for the calibration of the model fell on θ_{WP}, because it was the only parameter that we could not retrieve from literature or measured directly, its value was only roughly estimated based on the substrate datasheet, without experimental validation. In contrast, all other parameters were obtained from the literature or measured during the experiments, making θ_{WP} the most uncertain parameter of the model.

Regarding the parameter importance, θ_{WP} had a variable position across the rankings, but had a value ($\rho_{\theta_{WP}} = 0.38, 0.22, 0.7$ for P1, W3 and W1 respectively) of not more than one order of magnitude different from the value for C_ϵ, while the collinearity indices γ_P for the subset $P = \{C_\epsilon, \theta_{WP}\}$ were 9.47, 11.15, 11.5 and 12.95 for P3, P1, W3, W1 respectively. Therefore, according to ([19]), the parameter vector $\boldsymbol{\theta}_P = \begin{bmatrix} C_\epsilon & \theta_{WP} \end{bmatrix}^T$ could be considered for model calibration.

6. Model calibration

The calibration of the model has been carried out with a simplex Nelder Mead minimization algorithm, a derivative-free optimization algorithm used to minimize a function in multiple dimensions, assuming as the cost function:

$$J(\boldsymbol{\theta}_P) = \sum_{i=1}^{n} \left(y(t_i) - \tilde{y}(t_i, \boldsymbol{\theta})\right)^2 \tag{61}$$

and obtaining the following estimates $\hat{\boldsymbol{\theta}}_P = \begin{bmatrix} \hat{C}_\epsilon & \hat{\theta}_{WP} \end{bmatrix}^T$:

$$\hat{C}_\epsilon = 32.479 \times 10^{-6}\ \mathrm{g \cdot J^{-1}} \tag{62}$$

$$\hat{\theta}_{WP} = \begin{cases} 0.0878 & \text{(peat)} \\ 0.2251 & \text{(wood)} \end{cases} \tag{63}$$

The 95 % confidence intervals of the estimated parameters can be computed as:

$$\Delta\hat{\theta}_{P,k} = \left[\hat{\theta}_{P,k} - 1.96 \cdot \mathrm{SE}_{P,k}(\hat{\boldsymbol{\theta}}), \hat{\theta}_{P,k} + 1.96 \cdot \mathrm{SE}_{P,k}(\hat{\boldsymbol{\theta}})\right] \tag{64}$$

where $\mathrm{SE}_{P,k}(\hat{\boldsymbol{\theta}})$ is the standard error for parameter $\hat{\theta}_{P,k}$:

$$\mathrm{SE}_{P,k}(\hat{\boldsymbol{\theta}}) = \sqrt{\sigma_{j,j}(\hat{\boldsymbol{\theta}})} \tag{65}$$

with $\theta_{P,k} = \theta_j$ and $\sigma_{j,j}(\hat{\boldsymbol{\theta}})$ being the diagonal element of the covariance matrix $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$:

$$\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) = J(\hat{\boldsymbol{\theta}}_P)\,\mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}) \tag{66}$$

and $\mathbf{I}(\hat{\boldsymbol{\theta}})$ the Fisher information matrix. In turn, matrix $\mathbf{I}(\hat{\boldsymbol{\theta}})$ is computed as:

$$\mathbf{I}(\hat{\boldsymbol{\theta}}) = \mathbf{S}^T(\hat{\boldsymbol{\theta}})\,\mathbf{W}\,\mathbf{S}(\hat{\boldsymbol{\theta}}) \tag{67}$$

with $\mathbf{W}$ being a diagonal weighting matrix, whose elements are the inverse of the estimated variances of the measurements:

$$w_{i,i} = \frac{1}{\alpha^2\, y^2(t_i)} \tag{68}$$

where α is an estimate of the relative standard deviation of measurement errors, here assumed equal to 5 %. The confidence intervals for the estimated parameters are therefore:

$$\Delta_{\hat{C}_\epsilon} = \left[22.474 \times 10^{-6},\ 42.484 \times 10^{-6}\right]\ \mathrm{g \cdot J^{-1}} \tag{69}$$

$$\Delta_{\hat{\theta}_{WP}} = \begin{cases} [0.06135,\ 0.11425] & \text{(peat)} \\ [0.19865,\ 0.2515] & \text{(wood)} \end{cases} \tag{70}$$

The simulated dry weight after calibration for each treatment, compared to the measured and nominal simulated ones, are shown in Figs. 12, 13, 14, 15. The captions report the values of the Theil's inequality coefficient (TIC) ([11]) and the modified mean absolute relative error (MARE) coefficient, used to assess the goodness of fit:

$$\mathrm{TIC} = \frac{\sqrt{\sum_{i=1}^{n}\left(y(t_i) - \tilde{y}(t_i, \hat{\boldsymbol{\theta}})\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(y(t_i)\right)^2} + \sqrt{\sum_{i=1}^{n}\left(\tilde{y}(t_i, \hat{\boldsymbol{\theta}})\right)^2}} \tag{71}$$

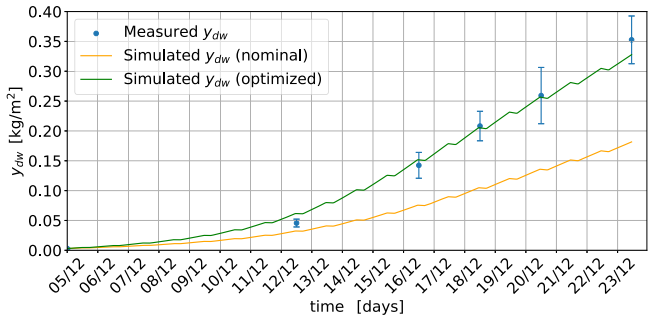

Fig. 12. P3: calibrated, nominal and measured dry weight TIC = 0.02754, MARE = 0.1165.

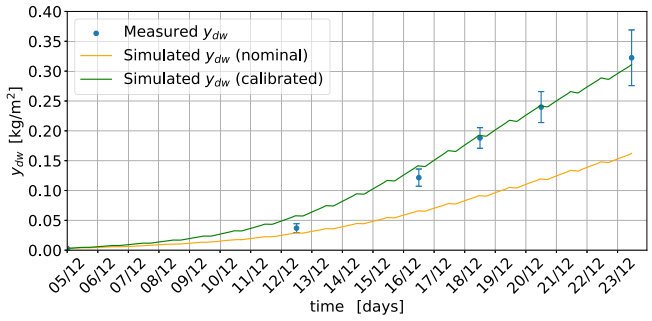

Fig. 13. P1: calibrated, nominal and measured dry weight TIC = 0.03454, MARE = 0.1252.

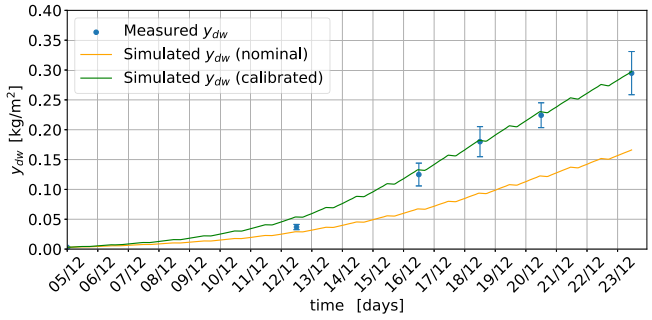

Fig. 14. W3: calibrated, nominal and measured dry weight TIC = 0.02269, MARE = 0.1066.

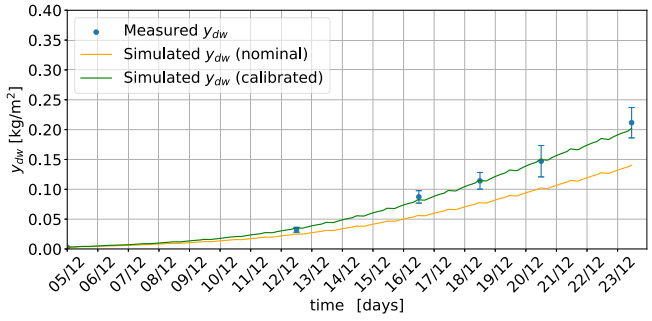

Fig. 15. W1: calibrated, nominal and measured dry weight TIC = 0.04403, MARE = 0.1413.

$$\text{MARE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y(t_i) - \tilde{y}(t_i, \hat{\boldsymbol{\theta}})\right|}{y(t_i) + \phi} \tag{72}$$

where the small correction factor ϕ (10^{-6}) is applied to avoid division by zero. For both indices, a closer value to zero indicates better performance.

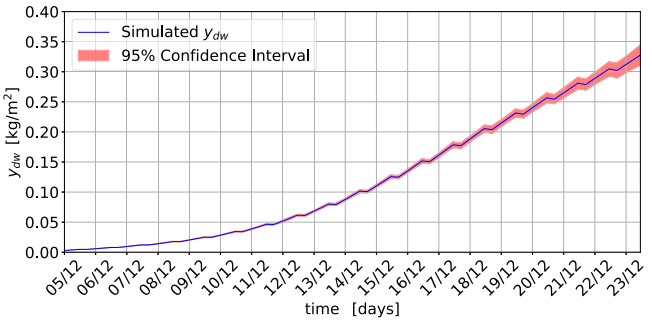

Fig. 16. P3: Dry weight at 95 % confidence interval.

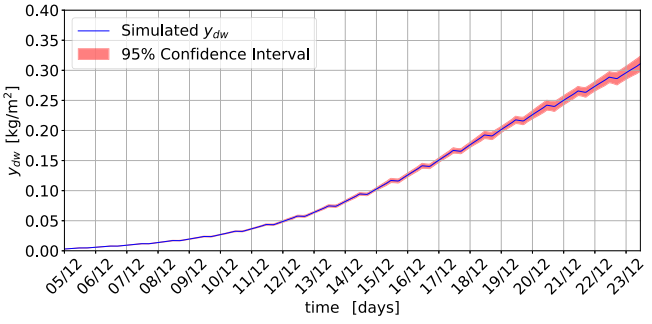

Fig. 17. P1: Dry weight at 95 % confidence interval.

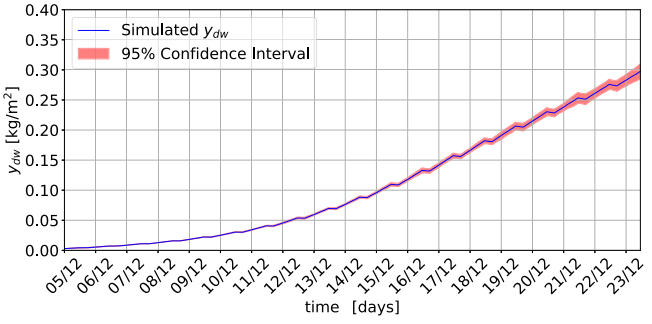

Fig. 18. W3: Dry weight at 95 % confidence interval.

Confidence intervals can be also estimated with reference to model predictions, taking into account the variability in parameter estimates:

$$\Delta\tilde{y}(t_i, \hat{\boldsymbol{\theta}}) = \left[\tilde{y}(t_i, \hat{\boldsymbol{\theta}}) - 1.96 \cdot \text{SE}_{y,i}(\hat{\boldsymbol{\theta}}), \tilde{y}(t_i, \hat{\boldsymbol{\theta}}) + 1.96 \cdot \text{SE}_{y,i}(\hat{\boldsymbol{\theta}})\right] \tag{73}$$

where the standard error $\text{SE}_{y,i}(\hat{\boldsymbol{\theta}})$ is defined as:

$$\text{SE}_{y,i}(\hat{\boldsymbol{\theta}}) = \sqrt{\mathbf{s}_i(\bar{\boldsymbol{\theta}})\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})\mathbf{s}_i^T(\bar{\boldsymbol{\theta}})} \tag{74}$$

with $\mathbf{s}_i(\bar{\boldsymbol{\theta}})$ denoting, with a slight abuse of notation, the i-th row of the sensitivity matrix $\mathbf{S}(\boldsymbol{\theta})$. Figs. 16, 17, 18, 19 provide a visual representation of the uncertainty associated with the model predictions, taking into account the variability in parameter estimates.

The dynamics of volumetric substrate water content θ across all four treatments, as illustrated in Fig. 20, yield finally some noteworthy observations. As expected, peat led to higher dry weight due to its better water retention properties. However, peat extraction not only releases large amounts of carbon dioxide but also disrupts fragile ecosystems that take thousands of years to regenerate. In contrast, wood fiber is typically produced from renewable forestry by-products that would otherwise be discarded as waste. However, Fig. 20 suggests that wood fiber can potentially serve as a viable alternative for more sustainable lettuce production, despite its lower water retention capacity. By increasing the frequency of daily irrigation events, substrate water content could be maintained near field capacity, effectively compensating for wood

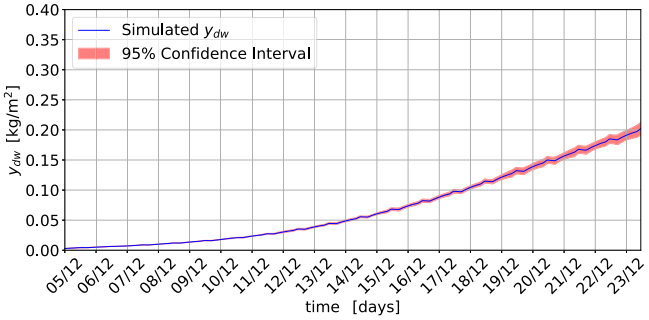

Fig. 19. W1: Dry weight at 95 % confidence interval.

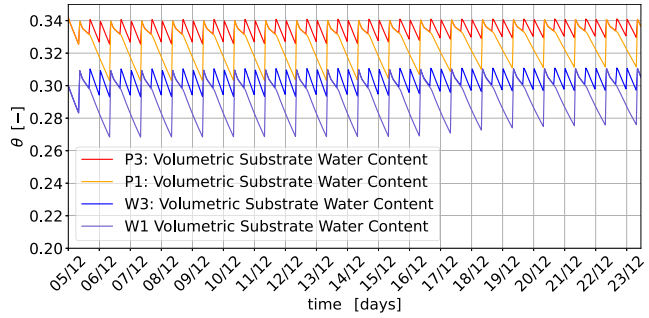

Fig. 20. Volumetric substrate water content of the four treatments.

fiber's limitations and achieving results comparable to peat-based cultivation.

7. Conclusion

In this paper, a dynamic growth model proposed for lettuce has been extended to integrate substrate water content dynamics, to account for the effects of water stress on biomass accumulation.

Most model parameters have been adopted from peer-reviewed studies or experimentally determined. Estimation methods and primary sources are reported in the text, while associated uncertainty for calibrated parameters is provided, based on sensitivity analysis and model fitting.

Under water-stressed conditions, the model initially underestimated dry weight production, which was primarily attributed to the superior light utilization efficiency of indoor farming compared to greenhouse environments where the model was originally designed. Following calibration of the light use efficiency and wilting point substrate water content parameters, the model achieved high precision in predicting dry weight production under all tested conditions. These findings demonstrate the model's value as a decision-support instrument for enhancing irrigation management and substrate selection in controlled environment agriculture.

The enhanced model offers flexible adaptation for investigating various substrate effects through parameter adjustments related to water retention characteristics. Additionally, it enables application to diverse soilless growing systems, including hydroponic and potentially aeroponic configurations. Such adaptations would lead to different representation of plant-substrate interactions, supporting the optimization of both yield and resources usage across varying setups. The model developed in this thesis provides a solid foundation for these extensions, offering a reliable representation of substrate water content dynamics.

CRediT authorship contribution statement

Davide Marino: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization; **Alessandro Antona:** Writing – original draft, Validation, Investigation, Data curation, Conceptualization; **Arianna Catenacci:** Writing – review & editing, Writing – original draft, Software, Methodology; **Gianni Ferretti:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Conceptualization.

Data availability

The authors do not have permission to share data.

Declaration of competing interest

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

Appendix A. Parameters of the Van Henten model

Table A.1
Parameter values used in the Van Henten (1994) lettuce growth model.

Parameter	Value	Units	Description
C_α	0.68	–	Factor converting assimilated CO_2 into sugar
C_β	0.80	–	Yield factor indicating respiratory and synthesis losses of non-structural material
$C_{gr,\max}$	5×10^{-6}	s^{-1}	Maximum specific growth rate at 20 °C, under saturating conditions
$C_{Q10,gr}$	1.6	–	Q_{10} factor for growth rate
$C_{resp,sht}$	3.47×10^{-7}	s^{-1}	Maintenance respiration rate coefficient for shoots
$C_{resp,rt}$	1.16×10^{-7}	s^{-1}	Maintenance respiration rate coefficient for roots
$C_{Q10,resp}$	2	–	Q_{10} factor for respiration
C_τ	0.15	–	Ratio of root dry weight to total crop dry weight
C_K	0.9	–	Extinction coefficient
C_{lar}	75×10^{-3}	$m^2 \cdot g^{-1}$	Structural leaf area ratio
C_ω	1.83×10^{-3}	$g \cdot m^{-3}$	Density of CO_2
C_Γ	40	ppm	CO_2 compensation point at 20 °C
$C_{Q10,\Gamma}$	2	–	Q_{10} factor for Γ
C_ϵ	17×10^{-6}	$g \cdot J^{-1}$	Light use efficiency
G_{bnd}	0.007	$m \cdot s^{-1}$	Boundary layer conductance
G_{stm}	0.005	$m \cdot s^{-1}$	Stomatal conductance
$C_{car,1}$	-1.32×10^{-5}	$m \cdot s^{-1} \cdot °C^{-2}$	Quadratic coefficient for carboxylation conductance
$C_{car,2}$	5.94×10^{-4}	$m \cdot s^{-1} \cdot °C^{-1}$	Linear coefficient for carboxylation conductance
$C_{car,3}$	-2.64×10^{-3}	$m \cdot s^{-1}$	Constant term for carboxylation conductance
N_{plants}	18	m^{-2}	Plant density

Appendix B. Parameters of the extended model

Table B.2
Parameter values used in the extended model (nominal).

Parameter	Value	Units	Description
γ	0.0672	$kPa \cdot °C^{-1}$	Psychrometric constant
α	0	–	Canopy reflection coefficient
$h_{c,\max}$	0.3	m	Maximum crop height
$K_{cb,TAB}$	0.9	–	Basal crop coefficient for lettuce in standard conditions
θ_{FC}	0.342 (peat) 0.312 (wood)	–	Substrate water content at field capacity
θ_{WP}	0.1756 (peat) 0.1693 (wood)	–	Substrate water content at wilting point
δ	0.06	$L \cdot s^{-1}$	Drainage rate
P	4.871×10^{-6} (peat) 1.013×10^{-6} (wood)	$L \cdot m^{-2} \cdot s^{-1}$	Percolation rate
L_f	0.075	m	Leaf characteristic dimension
Le	1.47	–	Lewis number for CO_2
K_d	20,000	$L^{-1} \cdot s$	Steepness of the transition for w_D
K	200	$W^{-1} \cdot m^2$	Steepness of the transition for $f_{\text{day/night}}$
a	n.d.	m	Width of a single pot
b	n.d.	m	Length of a single pot
z	n.d.	m	Depth of a single pot
η	10	$W \cdot m^{-2}$	PAR threshold
q	1.5	–	Curvature degree of $k_s - \theta$ relationship

Appendix C. Variables of the extended model

Table C.3
Variables of the model (Part 1).

Variable	Units	Description
State Variables		
x_{nsdw}	$g \cdot m^{-2}$	Non-structural dry weight
x_{sdw}	$g \cdot m^{-2}$	Structural dry weight
h	L	Water content in substrate
Growth and Photosynthesis Variables		
r_{gr}	s^{-1}	Growth rate
f_{resp}	$g \cdot m^{-2} \cdot s^{-1}$	Respiration flux
f_{phot}	$g \cdot m^{-2} \cdot s^{-1}$	Photosynthesis flux
$f_{phot,max}$	$g \cdot m^{-2} \cdot s^{-1}$	Maximum photosynthesis flux
Γ	ppm	CO_2 compensation point
ϵ	$g \cdot J^{-1}$	Light use efficiency
y_{dw}	$kg \cdot m^{-2}$	Total dry weight
Gas Exchange Variables		
G_{CO_2}	$m \cdot s^{-1}$	Total CO_2 conductance
G_{car}	$m \cdot s^{-1}$	Carboxylation conductance
G_{bnd}	$m \cdot s^{-1}$	Boundary layer conductance
Evapotranspiration Variables		
ET_0	$mm \cdot s^{-1}$	Reference evapotranspiration
Δ	$kPa \cdot °C^{-1}$	Slope of saturation vapor pressure curve
C_d	$s \cdot m^{-1}$	Aerodynamic resistance parameter
e	kPa	Saturation vapor pressure at air temperature
e_a	kPa	Actual vapor pressure
e_s	kPa	Mean saturation vapor pressure
R_n	$MJ \cdot m^{-2} \cdot s^{-1}$	Net radiation
G	$MJ \cdot m^{-2} \cdot s^{-1}$	Substrate heat flux

Table C.4
Variables of the model (Part 2).

Variable	Units	Description
Water Stress and Crop Coefficients		
K_e	–	Soil evaporation coefficient
K_r	–	Evaporation reduction coefficient
TEW	mm	Total evaporable water
REW	mm	Readily evaporable water
K_s	–	Water stress coefficient
K_{cb}	–	Basal crop coefficient
$K_{c,max}$	–	Maximum crop coefficient
f_c	–	Fraction of ground covered by vegetation
Water Balance Variables		
w_d	$L{\cdot}s^{-1}$	Drainage water flux
w_p	$L{\cdot}s^{-1}$	Percolation water flux
w_{et}	$L{\cdot}s^{-1}$	Evapotranspiration water flux
θ	–	Volumetric water content
Input Variables		
u_T	°C	Air temperature
u_{PAR}	$W{\cdot}m^{-2}$	Photosynthetically active radiation
u_{CO_2}	ppm	CO_2 concentration
u_2	$m{\cdot}s^{-1}$	Wind speed at 2 m height
u_{RH}	%	Relative humidity
u_{IRR}	$L{\cdot}s^{-1}$	Irrigation water input

References

[1] B. Alhnaity, S. Pearson, G. Leontidis, S. Kollias, Using Deep Learning to Predict Plant Growth and Yield in Greenhouse Environments, ArXiv:1907.00624, 2019.

[2] E.V. Henten, Validation of a dynamic lettuce growth model for greenhouse climate control, Agric. Syst. 45 (1) (1994) 55–72.

[3] J.L. Svensen, X. Cheng, S. Boersma, C. Sun, Chance-constrained stochastic mpc of greenhouse production systems with parametric uncertainty, Comput. Electron. Agric. 217 (2024).

[4] W. Rohde, F. Forni, Lettuce modelling for growth control in precision agriculture, Technical Report, arXiv preprint, 2023. Preprint.

[5] W. Sun, A. Coules, C. Zhao, C. Lu, A lettuce growth model responding to a broad range of greenhouse climates, Biosyst. Eng. 250 (2025) 285–305.

[6] P. Steduto, T.C. Hsiao, D. Raes, E. Fereres, AquaCrop–The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles, Agron. J. 101 (3) (2009) 426–437.

[7] D. Loconsole, G. Cocetta, P. Santoro, A. Ferrante, Optimization of led lighting and quality evaluation of romaine lettuce grown in an innovative indoor cultivation system, Sustainability 11 (3) (2019).

[8] M.F. López Mora, M.F. Quintero Castellanos, C.A. González Murillo, C. Borgovan, M. Salas Sanjuan, M. Guzmán, Predictive model to evaluate water and nutrient uptake in vertically grown lettuce under mediterranean greenhouse conditions, Horticulturae 10 (2) (2024).

[9] M.-H. Talbot, D. Monfet, Development of a crop growth model for the energy analysis of controlled agriculture environment spaces, Biosyst. Eng. 238 (2024) 38–50.

[10] R.G. Allen, L.S. Pereira, D. Raes, M. Smith, FAO Irrigation and Drainage paper No 56, Rome: Food and Agriculture Organization of the United Nations 56 (1998) 26–40.

[11] B. Decostere, J. Craene, S. Hoey, H. Vervaeren, I. Nopens, S.V. Hulle, Validation of a microalgal growth model accounting with inorganic carbon and nutrient kinetics for wastewater treatment, Chem. Eng. J. 285 (2016) 189–197.

[12] R.G. Allen, I.A. Walter, R.L. Elliott, T.A. Howell, D. Itenfisu, M.E. Jensen, R.L. Snyder, The ASCE Standardized Reference Evapotranspiration Equation, American Society of Civil Engineers, 2005.

[13] R.G. Allen, Skin layer evaporation to account for small precipitation events-An enhancement to the FAO-56 evaporation model, Agric. Water Manage. 99 (1) (2011) 8–18.

[14] A. Verhoef, G. Egea, Modeling plant transpiration under limited soil water: comparison of different plant and soil hydraulic parameterizations and preliminary implications for their use in land surface models, Agric. For. Meteorol. 191 (2014) 22–32.

[15] E.A. Ibrahim, N.E.S. Ebrahim, G.Z. Mohamed, Effect of water stress and foliar application of chitosan and glycine betaine on lettuce, Sci. Rep. 13 (1) (2023) 17274.

[16] J.W. Lee, J. Eom, W.H. Kang, J.H. Shin, J.E. Son, Prediction of transpiration rate of lettuces (lactuca sativa L.) in plant factory by Penman-Monteith Model, Protect. Hort. Plant Factory 22 (2013) 182–187.

[17] G. Egea, A. Verhoef, P.L. Vidale, Towards an improved and more flexible representation of water stress in coupled photosynthesis-stomatal conductance models, Agric. For. Meteorol. 151 (10) (2011) 1370–1384.

[18] T. Haslwanter, An Introduction to Statistics with Python: With Applications in the Life Sciences, Springer International Publishing, Cham, Cham, 2022.

[19] R. Brun, P. Reichert, H.R. Künsch, Practical identifiability analysis of large environmental simulation models, Water Resour. Res. 37 (4) (2001) 1015–1030.

[20] A. Catenacci, D. Carecci, A. Leva, A. Guerreschi, G. Ferretti, E. Ficara, Towards maximization of parameters identifiability: development of the CalOpt tool and its application to the anaerobic digestion model, Chem. Eng. J. 499 (2024) 155743.