DU4

Representing the FPSD, DU4 (Figure 0.1) focuses on the production of transgenic lettuce expressing bone hormone PTH-Fc in a Crop Growth Reactor (GCR). DU4 returns edible lettuce biomass, PTH-Fc mass, and waste (e.g. inedible biomass, transpired H_2O, spent media, gaseous oxygen) outputs from a CGR unit operation taking inputs of seed, growth media M, supply cargo (e.g. plant growth substrate, growth equipment), power (e^-) and photosynthetically available photons (γ). The CGR mathematics are governed by the modified energy cascade (MEC) model for crop growth. Media M is constructed from feedstocks of NO_3^-, CO_2, and H_2O.

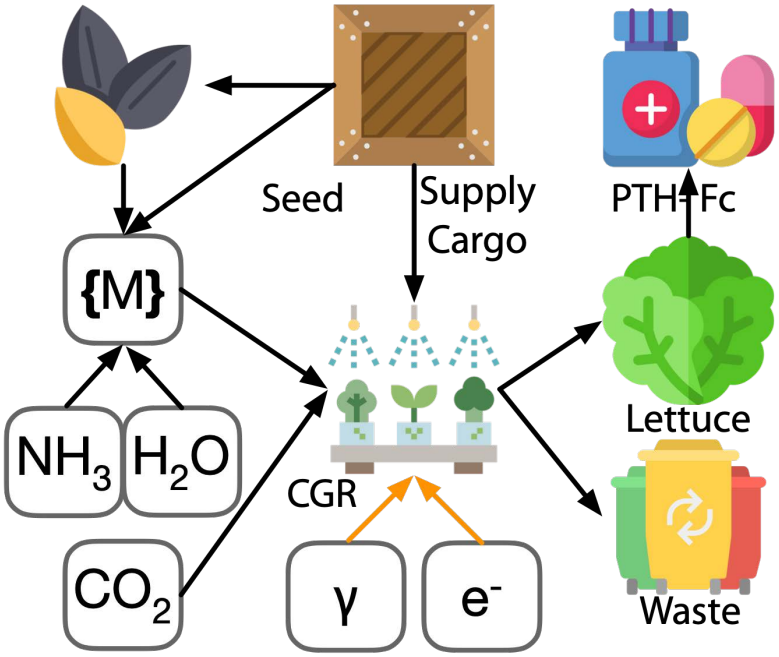

Figure 0.1: DU4

Model Description

We assume hydroponic growth for production of CM food in the form of $\mathcal{C} \in \{$Dry Bean, Lettuce, Peanut, Rice, Soybean, Sweet Potato, Tomato, Wheat, White Potato$\}$. In this particular case, the lettuce crop growth is modeled as a function of atmospheric concentration of carbon dioxide, $[CO_2]$, in $[\mu\text{mol}_{CO_2}/\text{mol}_{air}]$ and photosynthetic photon flux, PPF, in $[\mu\text{mol}_{photon}/\text{m}^2\cdot\text{s}]$. The output of the growth model is the crop growth rate, CGR, in $[\text{g}/\text{m}^2\cdot\text{d}]$. Integrating CGR provides total edible biomass (on a dry basis), TEB, in $[\text{g}/\text{m}^2]$, total crop biomass, TCB, in $[\text{g}/\text{m}^2]$.

We also propose to use this model to track the additional crop input requirements in terms of water and some set of nutrients provided at some initial start time $t_{0,i}$ of a hydroponic reactor j containing only 1 type of crop, $\mathcal{C}_i$. Also tracked will be the oxygen production.

At the end of a run both the waste, composed of reactor effluent, and the biomass B will harvested from some reactor j operating in the set of all reactors $\mathcal{J}(t)$. We propose a modeling framework that allows for the inclusion of testable variables affecting the biological process of growing a crop within a reactor and logistic variables such as number of reactors that contribute to the calculation of the ESM term. This is helpful when imagining a complete agricultural system composed of $n(t)$ reactors which may change in time in scenarios where reactors are produced or fail, defined as $\sum_j^{\mathcal{J}(t)} n_j(t)$ allowing us to represent the snapshot of all reactors with a specific crop i as $n_i(t) = \sum_i^{\mathcal{J}(t)} n_{i,j}(t)$.

We will model PTH-Fc production from a single reactor with an engineered lettuce crop simply as $\Xi = \Xi_{PTH-Fc} \cdot TEB_{let}$ where $PTH - Fc$ is the mass of the protein in [g], TEB_{let} is the total edible biomass of lettuce, and Ξ_{PTH-Fc} as the fraction of the protein contained in the edible portion of the lettuce plant in [g/g]. This approach allows for simple experimental evaluation of parameters with a single measurement of Ξ_{PTH-Fc}. We propose to use this simple model to calculate targets for PTH-Fc production for a mission given assumptions on the protein required by a crew-member under several conditions including (1) total lettuce diet; (2) only lettuce of a salad diet; (3) only lettuce of a complete diet; purified from a first lettuce crop not used for food. This work is on-going.

Model Mathematics

The following mathematical model was adopted for use in CUBES from literature on the modified energy cascade (MEC) model[1, 2] and from NASA references[3]. The use of this model within CUBES has been described to some extent beginning with the CUBES Y2Q2 report.

Biomass for a single plant k in a single reactor j of some crop i is denoted as $B_{i,j,k}$ and formulated as a continuous differential equation by

$$\frac{dB_{i,j,k}}{dt} = CGR_i(t, [CO_2], PPF) = \frac{MW_{C,i}}{BCF_i}\mathrm{DCG}_i(t, [CO_2], PPF) \tag{1}$$

$$= \frac{MW_{C,i}}{BCF_i}(0.0036) \cdot \mathrm{H}_i \cdot PPF_i. \cdot CUE_i(t, [CO_2])... \tag{2}$$

$$... \cdot CQY_i(t, [CO_2], PPF_i) \cdot A_i(t, [CO_2], PPF_i) \tag{3}$$

where CGR is crop growth rate in [g/(m^2·d)], MW_C is the molecular weight of Carbon in [g/mol], BCF is biomass carbon fraction [unitless], and DCG is the daily carbon gain in [mol$_C$/(m^2· d)]. The DCG can be represented in terms of functions

$$CQY = \begin{cases} CQY_{MAX} & t \le t_Q \\ CQY_{MAX} - (CQY_{MAX} - CQY_{MIN})\frac{t-t_Q}{t_m-t_Q} & t_Q < t \le t_M \end{cases} \tag{4}$$

$$CUE = \begin{cases} CUE_{MAX} & t \le t_Q \\ CUE_{MAX} - (CUE_{MAX} - CUE_{MIN})\frac{t-t_Q}{t_M-t_Q} & t_Q < t \le t_M \end{cases} \tag{5}$$

$$A = \begin{cases} A_{MAX}\left(\frac{t}{t_A}\right)^n & t < t_A \\ A_{MAX} & t \ge t_A \end{cases} \tag{6}$$

for the photo-period H in [h/d], 24-hour carbon use efficiency CUE [unitless], canopy quantum yield CQY in $[\mu\text{mol}_C/\mu\text{mol}_{PPF}]$, photosynthetic photon flux PPF in $[\mu\text{mol}_\gamma/(\text{m}^2\cdot\text{ s})]$, n is a crop-based exponent, A is the unitless fraction of PPF absorbed by the plant canopy, t is the time variable in [d] which is actually time after emergence of a crop and thus is measured in $[d_{AE}]$, t_M is the time at maturity in $[d_{AE}]$, t_Q is the time at onset of canopy senescence in $[d_{AE}]$, and t_A is the time until canopy closure in $[d_{AE}]$ which can be calculated as a function of the effective photosynthetic flux PPF_E, photo-period H, and the nominal photo-period per crop H_0 as

$$PPF_E = PPF\frac{H}{H_0} \tag{7}$$

$$t_A(PPE_E, [CO_2]) = \sum_{u\in[-1,3]}\sum_{v\in[-1,3]} D_{u,v}\cdot PPF_E^v\cdot[CO_2]^u \tag{8}$$

where $[CO_2]$ is the concentration of carbon dioxide in the reactor $[\mu\text{mol}_{CO_2}/\text{mol}_{air}]$. The coefficients $D_{i,j}$ for indices u and v in range [-1,3] are calculated experimentally in provided from tables for each crop in $\mathcal{C}$. The indices u and v are also used as exponents u and v within the double sum. Likewise, we calculate CQY_{MIN} and CQY_{MAX} in $[\mu\text{mol}_C/\mu\text{mol}_{PPF}]$, both as crop specific minimal and maximal values of CQY, with a similar method to t_A as

$$CQY_{MAX}(PPE, [CO_2]) = \sum_{u\in[-1,3]}\sum_{v\in[-1,3]} C_{u,v}\cdot PPF^v\cdot[CO_2]^u \tag{9}$$

for indices u and v in range [-1,3] are calculated experimentally in provided from tables for each crop in $\mathcal{C}$. The sets of 25 coefficients $C_{u,v}$ and $D_{u,v}$ are calculated per crop by multivariate polynomial regression.

We also begin to represent the remaining species as

$$\frac{dCO_2}{dt} = r_{CO_2} - DCG(t, PPF, [CO_2]) \tag{10}$$

$$\frac{dH_2O}{dt} = r_{H_2O} - DCG(t, PPF, [CO_2]) + DTR(t, PPF, [CO_2]) \tag{11}$$

$$\frac{dO_2}{dt} = -r_{O_2} + OPF\cdot DCG(t, PPF, [CO_2]) \tag{12}$$

where r_s corresponds to a rate of an added species s, DTR is the daily transpiration rate, and OPR is the oxygen production rate. DCG, DTR, and DCG are functions of time. We can model the daily transpiration rate DTR using additional equations as

$$DTR = 3600(H)\left(\frac{MW_W}{\rho_W}\right)g_C\left(\frac{VPD}{P_{atm}}\right) \tag{13}$$

where $MW_W = 18.015$ g/mol is the molecular weight of water, P_{atm} is the total atmospheric pressure in the growth environment, $\rho_W = 998.23$ [g/L] is the density of water (this will be temperature T dependent), and H is again the photoperiod. The g_C term in this

differential equation is the canopy surface conductance and is a function of the canopy stomatal conductance g_S and the atmospheric aerodynamic conductance g_A as

$$g_C = \frac{g_A g_S}{g_A + g_S} \tag{14}$$

$$g_S = \begin{cases} (1.717T_\gamma - 19.96 - 10.54VPD)\left(\frac{P_{net}}{[CO_2]}\right) & \text{planophile} \\ (0.1389T_\gamma - 15.32RH)\left(\frac{P_{net}}{[CO_2]}\right) & \text{erectophile} \end{cases} \tag{15}$$

$$g_A = \begin{cases} 2.5 & \text{planophile} \\ 5.5 & \text{erectophile} \end{cases} \tag{16}$$

where the units of g_A, g_C, g_S are [$\text{mol}_{water}/\text{m}^2 \cdot$ s], the constants are RH is the atmospheric relative humidity, T_γ is the temperature with the lights on and the non-constant term VPD is the vapor pressure deficit which can be calculated as

$$VPD = VP_{SAT} - VP_{AIR} \tag{17}$$

$$= 0.611e^{\frac{17.4T_\gamma}{T_\gamma + 239}}(1 - RH) \tag{18}$$

where VP_{SAT} and VP_{AIR} are the saturated vapor pressure and the actual atmospheric vapor pressure of the reactor in [kPA]. We can therefore determine the important terms of gross canopy photosynthesis P_{gross} and net canopy photosynthesis P_{net} as

$$P_{gross} = (A)(CQY)(PPF) \tag{19}$$

$$P_{net} = \left(\frac{D_{PG} - H}{D_{PG}} + \frac{(H)(CUE}{D_{PG}}\right) P_{gross} \tag{20}$$

where D_{PG} is the plant growth diurnal cycle in [h/d].

Model Implementation

While an initial CGR bioreactor model has been implemented within the echusOverlook framework, the changing structure of the echusOverlook codebase, described in the RMA section, had lead to complications in the stability for using the GCR model.

Additional Model Considerations

There are several considerations to adoption and validation of the modified energy cascade model as the representative mathematics for DU4. These considerations are

1. crop growth model exclusion of plant-made pharmaceutical production,
2. crop cultivar mismatch between data used for model parameterization and experimental selection, and
3. lighting system mismatch between data used for model parameterization and experimental selection.

Here we define each of these considerations, highlight opportunities, and chart a course for model adaptation, validation, and extension.

Plant-made pharmaceuticals

Consideration: The model does not account for *in planta* accumulation of total soluble proteins and high-value recombinant proteins.

Opportunity: Developing an extension of the modified energy cascade model to encompass total soluble protein and plant-made pharmaceuticals could provide unique insight into pharmaceutical production behavior and optimization and/or the impact of recombinant protein production on photosynthetic growth. There has been limited investigation of environmental condition design space in the production of plant-made pharmaceuticals. The model could provide guidance to experimental design for optimization.

Path Forward: Although our ultimate goal would be to test homozygous PTH-Fc transgenic lettuce lines, as a preliminary study we could test heterozygous transgenic lettuce lines (which we currently have available) for PTH-Fc expression levels (e.g. 10 mg PTH-Fc/kg fresh weight lettuce) and biomass production for inclusion in the Year 3 demo model. This preliminary study will seek to define PTH-Fc expression as a function of biomass and plant age. For the Year 5 demo, we would like to define PTH-Fc expression in homozygous transgenic lettuce lines as a function of biomass, plant age, atmospheric CO_2, and PPF.

Crop cultivar

Consideration: The model contains data generated for lettuce cultivar Waldmann's Green, while experimental transgenic lettuce cultivars used are Crisphead Great Lakes and Romaine Paris Island.

Opportunity: Generation of model parameters for additional lettuce cultivars Crisphead Great Lakes and Romaine Paris Island have the potential to expand the model database.

Path Forward: We acknowledge the cultivar mismatch and choose to use the model parameters obtained for Waldmann's Green in the Year 3 demo. When homozygous transgenic lettuce lines expressing PTH-Fc are established, growth of wild-type Crisphead Great Lakes and Romaine Paris Island cultivars will serve as control conditions for experimental testing. At this juncture, the value to re-visiting and closing the model mismatch is higher.

Lighting system

Consideration: The model data is generated using high-pressure sodium lighting and the model is not sensitive to light spectrum (or light recipes) as a variable input although we know that light spectrum has a significant influence on outputs such as the biomass growth rate. Existing experimental lighting systems are LED-based. There is limited scientific incentive in converting lighting systems to high-pressure sodium for experimentation.

Opportunity: Expansion of the model to include sensitivity to the spectrum of the light source would represent a substantial contribution to science. However, that task may or may not prove to be feasible within the constraints of CUBES. At minimum, incorporating initial insights would set the foundation for additional experimentation. Performing light spectrum experimentation with transgenic lettuce lines expressing PTH-Fc would also serve as a considerable contribution to plant-made pharmaceutical research.

Path Forward: We plan to leverage CUBES-generated experimental data on growth of Waldmann's Green to identify preliminary sensitivity of the model to 30% blue 70% red and

cool white (30% blue) LED light spectrums. Based on results of sensitivity analysis, we will either accept or reject sets of existing data for model validation.

Experimental Design

In the current proposed path forward, two sets of experiments have been identified for completion of DU4 for the Year 3 demo:

1. **Quantification of PTH-Fc expression level in T_1 heterozygous lettuce lines:** T_1 generation heterozygous lettuce lines will be screened for presence of PTH-Fc using PCR for gene detection and a dot blot assay for protein detection. We will collect biomass samples from five positive lines when the plants are approximately seven weeks of age post-seeding, which has been reported in literature to be an optimal harvest age for transgenic lettuce lines expressing β-glucuronidase (GUS) protein under control of the Cauliflower mosaic virus-35S promoter[4]. We will use the colorimetric Bradford assay to measure total soluble protein content in these samples and an enzyme-linked immunosorbent assay (ELISA) to quantify the mass of PTH-Fc in the harvested plant tissue (mg PTH-Fc/kg FW lettuce) of these samples.

2. **Sensitivity analysis of model output biomass to spectrum of the light source:** We will be performing sensitivity analysis of model output biomass to spectrum of the light source using CUBES-generated data on Waldmann's Green lettuce growth under 30% blue 70% red and cool white (30% blue) LED light spectrums with a fixed PPF of 300 $\mu mol_{photon}/m^2 \cdot s$ and atmospheric CO_2 concentration of 800 ppm. The variation of the biomass output in the different light spectrum conditions will be calculated and assessed for acceptability to meeting model needs.

1 References

[1] Harry Jones and James Cavazzoni. *Top-level crop models for advanced life support analysis*. Tech. rep. 2000.

[2] J Cavazzoni. "Using explanatory crop models to develop simple tools for Advanced Life Support system studies". In: *Advances in Space Research* 34.7 (2004), pp. 1528–1538. ISSN: 0273-1177. DOI: https://doi.org/10.1016/j.asr.2003.02.073. URL: http://www.sciencedirect.com/science/article/pii/S0273117704006192.

[3] Molly S Anderson, Michael K Ewert, and John F Keener. "Life support baseline values and assumptions document". In: (2018).

[4] Tsuyoshi Okayama, Kenichi Okamura, and Haruhiko Murase. "Time Course of the expression of the CaMV35S-GUS gene in transgenic lettuce plants grown in a plant factory". In: *Engineering in Agriculture, Environment and Food* 2.3 (2009), pp. 83–88. ISSN: 1881-8366.