THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF CHEMICAL ENGINEERING

Development of an Analytic Method to Determine Algae Potential as Industrial Feedstocks for Biofuel Production

HINKAL PATEL
SUMMER 2021

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Chemical Engineering
with honors in Chemical Engineering

Reviewed and approved* by the following:

Wayne Curtis
Professor of Chemical Engineering
Thesis Supervisor

Michael Janik
Professor of Chemical Engineering
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

To determine the best algae feedstock for industrial hydrocarbon production purposes, an intrinsic parameter to compare the productivity of algae species with varying growth rates, biomass compositions, and biosynthetic capabilities must be determined. This work proposes that hydrocarbon productivity, which is the rate of hydrocarbon production per unit process volume or area, can be used as a comparative parameter since it gives a gauge of the hydrocarbon production capabilities of any algae species under different growth conditions. Of particular interest is the isoprene hydrocarbon producing algae Botryococcus. One definition of hydrocarbon production in terms of hydrocarbon concentration produced per biomass per absorbed photons would be a particularly useful intrinsic parameter for a species of algae. This work also defines how to obtain biomass and hydrocarbon data from literature using digitization software and then how to calculate productivity from that data on the basis of light given to the algae as well as the amount of algae biomass within the culture vessel to yield an intrinsic productivity measure that will permit a meaningful comparison of different strains under different growth conditions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Curtis for all his help and mentoring throughout my undergraduate years. Research within his lab made up a significant portion of my four years at Penn State and I will continue to be thankful for all the information I've learned there through the rest of my career and beyond. Additionally, I would like to thank all of the lab members within Curtis Lab who are always willing to answer my questions and take the time to teach me new skills and concepts. I would like to thank Lucas Nugent and Noah Willis for teaching me the basis of algae research within the lab and supporting my own research. I'd also like to thank the current algae team which acted as a sounding board throughout my thesis-writing process and gave copious amounts of feedback. I would like to thank Dr. Janik for his support as my honors advisor; his advice throughout my college career has helped me greatly. Finally, I would like to thank my family for their continued support towards my education and my sister for her words of encouragement whenever I struggled to finish this thesis.

Key Terms and Abbreviations

Photosynthetically Active Radiation (PAR): Wavelengths of light that are utilized within photosynthesis, between 400 - 700 nm

Photosynthetic Photon Flux Radiation (PPFD): Rate of incident photons within the PAR range hitting a given area

Biologically Accessible Photons (BAP): Incident photons within the PAR range that are absorbed by the photosynthetic species

Hydrocarbon(s): HC

Average Specific Hydrocarbon Productivity: ASHP [g HC/g algae biomass/ (time or photon flux)

Average Volumetric Hydrocarbon Productivity: AVHP [g HC/culture volume/time or photon flux)

Instantaneous Volumetric Hydrocarbon Productivity: IVHP

Chapter 1

Background

In the last two centuries, carbon dioxide emissions rose from 49.95 million tons in 1819 to 3.02 billion tons in 1919 and continue to rise past the 36.42 billion tons emitted in 2019.[9] As a result, the carbon dioxide concentration in the atmosphere has risen from fluctuating around 225 ppm for the last 600,000 years to past 400 ppm in the last few years.[9] The unprecedented rise in carbon dioxide levels in the atmosphere has been linked to many changes in climate, such as increasing temperatures, rising sea levels, and more severe weather natural disasters. Paired with the economic and physical limits of fossil fuels and the increased demand for energy due to rising global population, a majority of nations have vowed to achieve net-zero carbon dioxide emissions by mid-century as well as invest in sustainable energy sources, such as biofuels. Biofuels are cleaner than fossil fuels since they have 10-45% oxygen levels and little sulfur emission.[10] Currently, most biofuels are produced from food crops such as corn, sugarcane, wheat, and sugarbeet, raising the question of how to partition the crops between food and energy production.[10] A possible solution for sustainable energy production is algae biofuels, which can consume carbon dioxide to photosynthetically produce biofuels. However, the photoautotrophic nature of algae requires both a carbon dioxide and water source for growth and biofuel production. Additional research is also necessary to increase the efficiency of algal input to chemical energy production, optimizing hydrocarbon extraction from algae cells, and determining the best algae feedstock for industrial biofuels production to make algae biofuels a

more economically viable competitor for fossil fuels in addition to its role in CO2 emission reduction.

Decades of research has targeted the issue of feedstock selection by conducting biomass growth and fuel feedstock production analyses of microalgae species such as *Botryococcus braunii, Chlorella vulgaris,* and *Chlamydomonas reinhardtii.* Microalgae have been grown under varying light intensity, growth mediums, pH, temperature, carbon dioxide supplementation, and reactors in order to optimize hydrocarbon production. However, the diverse experimental approaches and lack of consistent analysis has resulted in heterogeneous reports of algae growth and hydrocarbon production that cannot be compared to other algae growth literature. In order to determine the algae species best suited for industrial biofuel production, there is a need for a comparative parameter that can be used to compare hydrocarbon production by different algae species grown under different conditions and within different reactors or in open-culture systems. This thesis presents biomass productivity as an intrinsic parameter that can compare the hydrocarbon production of algae under any growth conditions and processing methods for hydrocarbon extraction.

Prior algae growth research within Curtis Lab has compared photosynthetic algae and bacteria to determine the optimal feedstock for biofuels production. Theses by Lisa Grady and Amalie Tuerk in 2010 and 2011 respectively have shown that although the doubling time for the microalgae *Botryococcus braunii* is thirteen times longer than *Chlorella vulgaris, B. braunii* is able to store 2.2 times more light energy into algae biomass than *C. vulgaris* and the two algae only had a 10% difference in biomass production per day, which shows that algae growth rate is not a good indicator of biomass production and energy conversion efficiency.[1,12] *B. braunii* is of particular interest since it can produce up to 80% of lipid content within its biomass and research

has shown between 90-97% hydrocarbon recovery with the Showa strain of *B. braunii* using n-hexane.[10, 13] For these reasons, all the calculations within the thesis will be done on experiments measuring the growth of the various strains of *B. braunii*, although the logic can be applied to any photoautotrophic algae species.

Chapter 2

Productivity as an Intrinsic Comparable Parameter

With the growing interest in microalgae as a promising alternative to fossil fuels, the question of which photosynthetic feedstock to use to produce biofuel arises. Production platform alternatives include microalgae, such as *Botryococcus braunii* and *Chlorella vulgaris*, or cyanobacteria, which all possess varying growth rates, biomass compositions, and biosynthetic capabilities. Given these fundamental differences, it is difficult to define a singular property to compare their potentials as a feedstock for biofuel production. There are intrinsic properties that reflect the biological rate-lmited potential of an organism, and there are operational performance parameters that reflect the ability of that organism to grown in a bioreactor which reflect process-limited potential, where light limitation is most relevant.

To determine the best candidate for industrial growth of algae, two criteria should be considered: the rate at which the algae species can produce biomass and the efficiency with which they produce the biomass using the nutrients and light they are given. Both criteria should be considered in tandem to dispel the lure of a fast-growing algae species that raises operating costs due to poor efficiency or a high-efficiency, slow-growing algae that cannot produce enough biomass within a certain period of time to cover capital and operating costs.

Biomass Productivity

The first criteria, rate of biomass production, corresponds to the productivity of the algae. Biomass productivity is defined in various permutations that relates the rate of biomass production per unit process volume (e.g. area, volume, cell).

The current emphasis on intrinsic growth rate in algae research means that most literature will use the experimental growth, such as the concentration of algae or accumulated biomass, plotted against time to calculate the productivity. This gives rise to two types of productivity calculations: average and instantaneous productivity. The average productivity is found when the overall concentration is divided by the length of time needed to reach that concentration as shown in **Equation 1**.

$$P_{avg} = \frac{\Delta X}{\Delta t} \qquad \textbf{Equation 1}$$

In comparison, instantaneous productivity is found from a tangent line to the growth curve at any specific time. This is represented by the derivative at a specific time (t_i):

$$P_{inst} = \left.\frac{dX}{dt}\right|_{t_i} \qquad \textbf{Equation 2}$$

Due to differing growth rates during the different stages of algae growth, as well as different efficiencies of hydrocarbon production at different culture densities, instantaneous productivity provides a basis for productivity that is more intrinsic to the culture, particularly when this is expressed per unit mass, or specific instantaneous productivity,

$$P_{specific,inst} = \frac{1}{X_{0_i}} \left.\frac{dX}{dt}\right|_{t_i} \qquad \textbf{Equation 3}$$

where X_{oi} is the total biomass concentration at the specific time (t_i) when the derivative is taken.

In contrast, average productivity is more closely related to the production output of the system, as it represents what has been produced over a functional operational period. Past work has shown that *B. braunii* and *C. vulgaris* differ in the time of their lipid production, in which *B. braunii* produces lipids in proportion to the concentration of algal biomass whereas *C. vulgaris* produces the majority of its lipids for storage after the algal biomass concentration has stopped increasing.[1] Therefore, calculating the average lipid productivity for *C. vulgaris* over the entire

growth and accumulation phases gives a very different meaning if compared to production such as *B. braunii* which accumulates a relatively constant yield. Additionally, the instantaneous productivity of algae is more relevant to predict growth within a continuous industrial algae growth system.

Another common basis for expressing algae productivity is based on per unit surface area, or areal productivity. This basis for productivity reflects the limitations of obtaining natural light, so that one can project to the 'footprint' of the algae growth systems. Concentration of algae (X) combine with volume to provide the absolute production (XV) that when combined with areal productivity characterizes a given photobioreactor system:

$$P_{areal} = \frac{1}{A}\frac{d(XV)}{dt} \qquad \textbf{Equation 4}$$

Since the volume of liquid being handled is related to operating costs, it is also typical to conduct some calculations of productivity based on volume. Volumetric productivity is related to areal productivity calculation based on the depth of culture, and simply:

$$P_{vol} = \frac{1}{V}\frac{d(XV)}{dt} \qquad \textbf{Equation 5}$$

which for a constant volume system is simply the change in algae concentration with time. Note that while depth provides greater scaled up volume, this does not translate to proportional increase in production due to the limitation of light availability. Since one of the main costs of industrial algae cultivation is replacement of growth medium in continuous-fed processes, so understanding the productivity of extracted volume provides insight into future operating costs.[3]

However, variations in the number of cells in a given area or volume can alter areal and volumetric productivities. If two identical reactors are under the same conditions, but one reactor contains more cells, then that reactor will likely produce more hydrocarbons. In order to obtain a

more intrinsic biological productivity for an algae species, divide the concentration of hydrocarbon produced (XV) by the concentration of algae:

$$P_{specific} = \frac{1}{X}\frac{d(XV)}{dt} \qquad \textbf{Equation 6}$$

The time component of the biomass productivity equation that is not typically presented in a comparable manner. In most literature, the algal growth curves are presented against the total culture time in days, which is not a comparable value since most research has the reactors under different light/dark cycles so the algae is not continuously exposed to high intensity light that can cause photoinhibition which damages the algae cells.[5] Since high density algae cultures are light-limited and their productivity is directly linked to the amount of light available, time should be measured in terms of photohours, which is the amount of time that the algae culture is supplied with light. Only counting periods of light helps normalize the various light/dark cycles cultures are grown under since the algae under dark periods are limited to the light-independent reactions of photosynthesis while continuously lighted systems do not have this limitation.

Calculating productivity based on the rate of photon use provides the kinetic information for scaling within a bioreactor system. At high cell concentrations, the optical density rapidly attenuates light, therefore, one can anticipate that any viable algae production system will become light-limited. Therefore, growth can be expressed as biomass produced (XV) per yield on photons absorbed from the photosynthetically active radiation (PAR) range, which is photosynthetic photon flux density (PPFD). Yield on a photon basis is then multiplied by PAR and lighted area in order to obtain productivity as grams of biomass produced per time as a basis of photons.

$$P_{photon} = \left(\frac{d(XV)}{photon}\right) * PAR * A = yield\ per\ photon * PAR * A \qquad \textbf{Equation 7}$$

In continuous algae growth systems, biomass and lipids are removed at a constant rate, called the dilution rate (D), throughout the experiment growth period and growth medium is replaced to provide nutrients to the cells in order to spur new biomass production to replace the biomass removed to extract hydrocarbons. Tuerk (2011) demonstrated that the net specific growth rate of algae under continuous, steady-state conditions is equal to the dilution rate of the system and under inorganic nutrient limiting conditions, the biomass productivity is dependent on the dilution rate and biomass concentration[1]:

$$P_{inorganic,nutrient-limited,continuous} = D * X \quad \textbf{Equation 8}$$

Algae cultures reach a steady state value in continuous systems and at steady state, light-limited conditions, the biomass concentration is dependent on the constant rate of overall culture increase (K), the dilution rate, and volume[1]:

$$X_{steady\ state} = \frac{K}{V*D} \quad \textbf{Equation 9}$$

Combining **Equation 8** and **Equation 9** yields the light-limited productivity of a continuous system:

$$X_{light-limited,continuous} = \frac{K}{V*D}(D * X) = \frac{K}{V} \quad \textbf{Equation 10}$$

All of the productivity definitions introduced above can also be applied to the production of specific biological components by the algae. For example, since the main concern of industrial algae growth processes is the production of hydrocarbons for use as biofuel, productivity can also be defined as the rate of production hydrocarbons over time by algae, which gives a

comparison between different algae species and strains in terms of hydrogen production based on reactor volume, area, or on a basis of biomass concentration.

Photosynthetic Efficiency

The second criteria of importance, efficiency of growth, is the percentage of solar input that is converted into biomass by an algae species and allows for the comparison of the ability of algae under different systems to effectively convert light into chemical energy. Photosynthetic efficiency can be calculated for total biomass produced, which includes all organic compounds such as algal biomass and oils, as well as other macronutrients such as carbohydrates, as well as just the efficiency of solar energy to oil.

The equation for light energy to oil energy efficiency has been defined in other literature[6] and compiled within the thesis by Tuerk (2011):

$$\eta_{oil} = \frac{P_{oil} * \Delta H_{oil}}{PAI * A} V \qquad \textbf{Equation 11}$$

where P_{oil} is the productivity of oil (g/L/photohour), ΔH_{oil} is the heat of combustion of the oil or hydrocarbon produced, PAI is photosynthetic active irradiance (W/m^2), A is lighted reactor area, and V is total culture volume within in the reactor.[1] The PAI corresponds to the light input energy, which can be calculated from the light intensity that the algae cultures are grown under. The light intensity corresponds to the incident PPFD. Although the total energy available from PPFD varies with the wavelength of light, all the photons are functionally equivalent in their ability to be biologically utilized for the excitation of photosystem electrons, so an average energy per mole quanta for light underwater can be used to convert from PPFD to PAI.[1] A more

rigorous process of determining the photon energy input for inconsistent lighting, such as sunlight, based on PAR photon flux is outlined within Chapter 3.

Chapter 3

Methods of Productivity Calculation

This chapter sets forth methods of calculating the various productivities defined in the previous chapter from a given biomass growth curve or converting between productivity definitions. However, not all the productivities defined in Chapter 2 can be calculated for within a given algae growth experiment depending on what information is provided within the paper. **Figure 1** shows the general process of algae growth and hydrocarbon extraction and what data both yield.

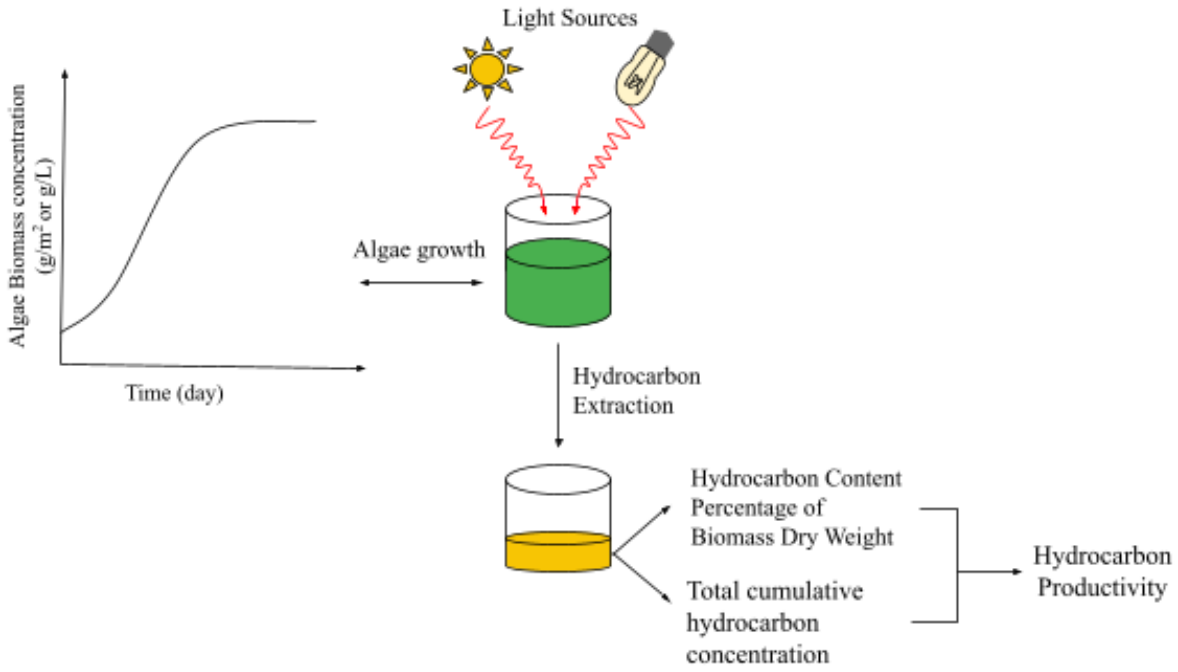


Figure 1. Process of hydrocarbon production and hydrocarbon productivity calculations

Determination of Algae Biomass Accumulation over Time

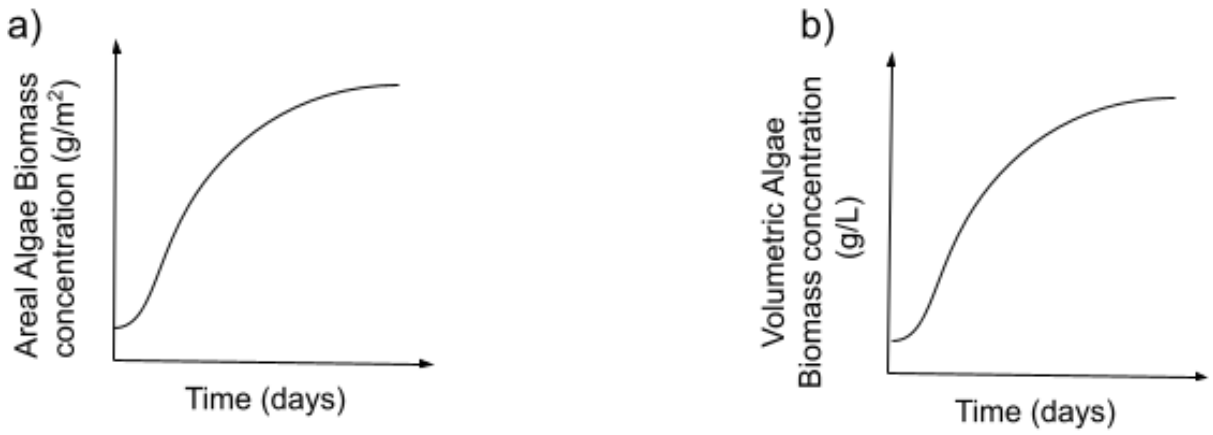


Figure 2. Algae biomass growth curves per areal (a) and volumetric (b) unit surfaces.

Research on algae growth experiments will typically have either algae biomass time course or tabulated data showing the concentration of algae biomass with respect to culture duration. Examples of algae biomass growth curves are shown in **Figure 2**. The biomass concentrations are usually given in terms of algal dry weight, but can also be expressed by cell count, optical density (OD), or chlorophyll fluorescence or content, as shown in **Table 1**. Many papers provide an equation or conversion factor to convert between optical density and algae dry weight, but the conversion factor varies based on algae species and strain as well as the wavelength used for the OD measurement. For *B. braunii* in particular, the production of extracellular hydrocarbons alters the inherent light absorption and scattering when OD samples are taken since the light must travel through the hydrocarbons before it is absorbed or scattered by the algae. For example, for the Race B Showa strain of *Botryococcus braunii* at a robust wavelength of 550 nm has an optical density-dry weight conversion of 2.15 ± 3 grams dry weight per robust OD.[4] The dry weight conversion for a *B. braunii* Race A strain cannot be determined using that value since Race A strains produce alkadienes and alkatrienes which have

differing light scattering and absorbance characteristics than the triterpenes produced by Race B strains. For Race A *B. braunii* strain Kützing, dry weight is reported to be calculatable by $2.6839*OD_{680} + 0.0987$ at a wavelength of 680.[15] However, the nature of light scattering is highly dependent on the size of the aggregates, and the Curtis laboratory has shown that aggregate size can change dramatically depending on the growth conditions, where large aggregates have a substantially reduced OD. Therefore, although OD is a convenient, quick way to obtain biomass information, it is not well suited for productivity comparisons under different conditions.

Table 1. Methods of Determining Algae Biomass Concentration

Method	Advantages	Disadvantages
Dry Weight (oven-drying)	Provide direct biomass values, convention for biomass concentration determination	Time consuming, tedious to heat samples to a constant weight
Dry Weight (freeze-drying)	Provide direct biomass values, reliable and efficient	Time-consuming, expensive
Optical Density	Rapid results, little algae volume is needed	Hydrocarbons and other materials in sample may alter the light absorbance and scattering, yielding an inaccurate OD for biomass concentration estimation
Chlorophyll Fluorescence or Content	Provides a good estimate of photosynthesis rate[11]	Varying cell sizes yield an inaccurate conversion from chlorophyll content to dry weight biomass[11]
Cell Count	Inexpensive[11]	Colonies and filament microalgae forms may be difficult to count[11]

If the growth data is not tabulated, software programs such as *Engauge Digitizer* can tabulate graphical data for calculation purposes.[7] The resulting data is dry weight biomass

concentration with respect to time, which is usually the total growth period of the algae during the experiment.

Biomass Volumetric and Areal Productivity

Volumetric and areal productivity can be determined directly from the dry weight biomass concentration time course data. If the biomass dry weight data is given in terms of mass per volume, then dividing the total change in biomass by the growth period, as shown in **Equation 1**, will yield the average volumetric biomass productivity. Likewise, if the biomass concentration is in terms of mass per area, then using **Equation 1** will yield the average areal biomass productivity. The average volumetric or areal productivity is the slope of the line connecting the initial and final condition as shown by the blue line in **Figure 3**. Average productivities should be calculated based on new biomass production over the growth period so the biomass introduced through the inoculation of the photobioreactor or remaining when biomass is recycled during an experiment does not artificially increase the biomass productivity. Both the average volumetric and areal biomass productivities are good indicators for the total biomass production under the given experimental growth conditions and the reactor utilization and performance.

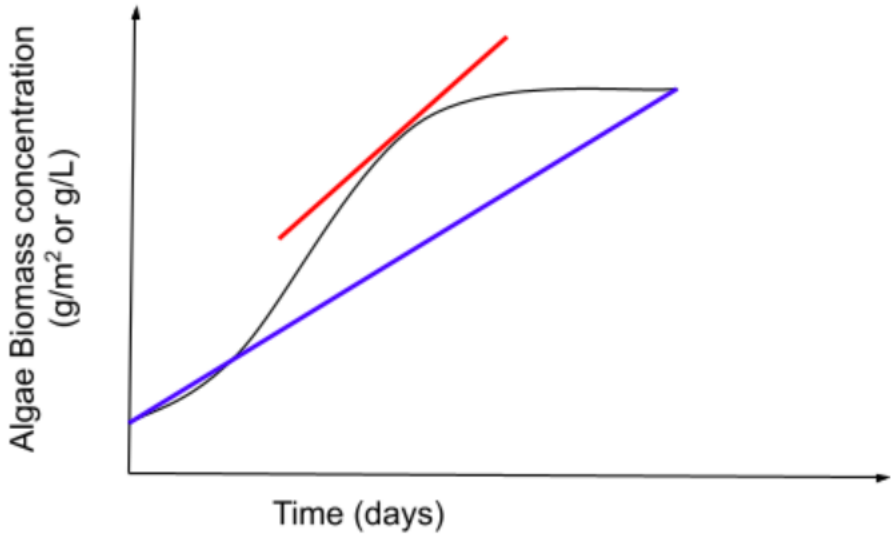


Figure 3. Average (slope of line connecting initial and final condition) versus instantaneous (slope of red tangent line) biomass productivity calculations from a biomass growth time curve

Determining Derivatives of Growth Data

To determine the derivative of growth data, the data must have enough data points to determine the slope of a sufficiently small region of data to use as the slope of the tangent line. An analytical approach to determine the slope of the tangent is to fit the experimental data to a trendline or an equation with adjustable parameters, such as the logistic function, which typically has four or five parameters. The logistic function is a sigmoid curve that would represent the initial lag period, exponential phase, the declining growth phase, and stationary phase well, as shown in **Figure 4**.

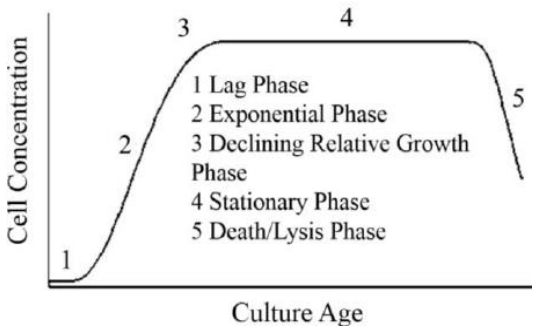


Figure 4. Algae growth phases with respect to algae culture age adapted from Farag & Price (2013)[16].

However, as soon as any equation is fitted to experimental data, the constraints associated with the equation become relevant when the derivative is taken, which makes traditional curve fitting less accurate for experimental growth data. This thesis instead proposes visual fitting of curves and the generating data points along the manually drawn curve in Chapter 4.

Conversion to Hydrocarbon Production per Volume or Area

As noted in Chapter 2, productivity can also be defined in terms of the rate of hydrocarbon production over time. In order to compare algae species as feedstocks for hydrocarbon production, the rest of this chapter will outline the steps to calculate algae hydrocarbon productivity as a basis of comparison between different reports of growth and production. Most literature papers present accumulating hydrocarbon concentration during the algae growth experiment determined after hydrocarbon extraction and calculate the average hydrocarbon volumetric or areal productivity over the total growth period. Similar to average biomass productivity, average hydrocarbon volumetric or areal productivity can indicate total production capability of an algae species under the experimental growth conditions and reactor set-up, but that requires an estimate of the hydrocarbon concentration within the reactor before

extraction. Although there are both physical (gravimetric) and chemical (analytical) methods of hydrocarbon production from algae, solvent extraction-based weights are often used in algae productivity experiments. Based on the theoretical hydrocarbon content of algae biomass and the hydrocarbon extraction yield, the hydrocarbon recovery percent can be calculated. For *B. braunii* strains, the hydrocarbon recovery percentage varies from 12% to 97% using 1,8-diazabicyclo-[5.4.0]-undec-7-ene/ethanol (1:1) and n-hexane respectively.[13] The hydrocarbon recovery of other solvents such as n-octane lie between the hydrocarbon recovery percentages of 1,8-diazabicyclo-[5.4.0]-undec-7-ene/ethanol (1:1) and n-hexane.[13] The range in hydrocarbon recovery shows that productivity calculations using the final extracted hydrocarbon concentration may not be indicative of the algae's actual hydrocarbon production capability since some hydrocarbons remain within the algae biomass. Notably, if the solvent extracts a breadth of hydrophobic algae metabolites, then a gravimetric assessed productivity will often be much higher than analytical methods that are based on measurement of very specific chemical species.

If the hydrocarbon recovery or extraction efficiency is given, the hydrocarbon yield after extraction can be corrected to determine the actual concentration of hydrocarbons produced by the algae:

$$X_{HC,produced} = \frac{X_{HC,extracted}}{Hydrocarbon\ recovery\ fraction} \qquad \textbf{Equation 12}$$

where $X_{HC,produced}$ is the hydrocarbon concentration within the reactor and $X_{HC,extracted}$ is the concentration of hydrocarbon after the extraction, which can be based on the solvent volume used to extract the hydrocarbons rather than the culture volume. It is important to have a precise understanding of the units of analysis, as often a researcher's presentation is intended to display trends, not be the basis of subsequent productivity calculations.

Hydrocarbon content may also be noted as a percentage of the biomass dry weight, which is determined through hydrocarbon extraction. Therefore, based on the extraction method listed within the literature, **Equation 12** can be used to correct hydrocarbon percentage of dry weight. The hydrocarbon concentration within the reactor can be calculated with the corrected hydrocarbon percentage and the biomass concentration in the reactor at the time when the hydrocarbon content was determined, which yields hydrocarbon concentration in terms of reactor volume or area.

Hydrocarbon volumetric and areal productivity is calculated in the same manner as biomass volumetric and areal productivity as described in the previous section using **Equation 4** or **Equation 5**. Again, only new hydrocarbon production should be used within the calculation to more accurately determine productivity over a certain period of time. Therefore, if hydrocarbon content is measured several times over the algae growth period, then the productivity over each sampling period should be based on the change in hydrocarbon concentration within the sampling period, not cumulative from the start of the experiment. Usually only average hydrocarbon volumetric productivity can be calculated if hydrocarbon concentration is only determined at the end of the experiment, as shown by the green area within **Figure 5** that shows the average volumetric productivity if only the final hydrocarbon concentration is given in the literature, which is represented by the green data point. Instantaneous productivity can be determined if the hydrocarbon concentration is determined multiple times over the algae growth period, as shown by the golden tangent lines in **Figure 5** which represent the varying instantaneous productivity within the reactor over time.

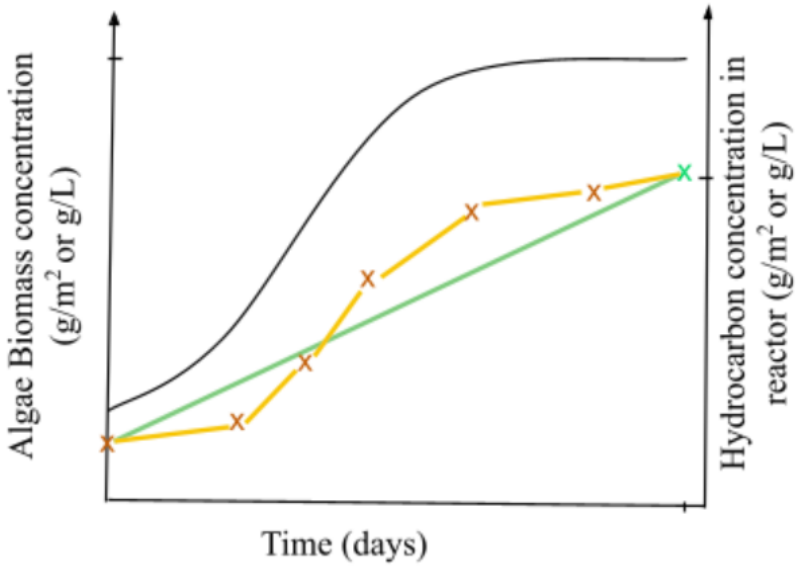


Figure 5. Hydrocarbon concentration within the reactor over the biomass growth period. The slopes over shorter periods are indicative the variations in instantaneous productivity that are experience throughout a batch culture period.

Figure 5 also shows that the average hydrocarbon productivity varies based on what time the biomass is sampled due to the different hydrocarbon production rates during the different algae growth phases. The average productivities during each theoretical sampling period is shown with the orange slopes, which only consider new hydrocarbon growth since the beginning of each sampling period. Summing the average productivity during each sampling period would yield an overall average hydrocarbon productivity that is more informative than the cumulative average productivity based on only the final green data point.

For continuous, light-limited systems, volumetric and areal are calculated using **Equation 8**, in which hydrocarbon concentration replaces the biomass concentration.

Conversion to Hydrocarbon Production per Dry Weight Biomass

Both areal and volumetric productivity fall short as comparative parameters for algae hydrocarbon production since they are dependent on the volume or surface area of the growth vessel and number of cells within that defined volume, meaning that a less productive algae species may appear to be as productive as an algae species that is intrinsically more productive if there are more cells within a given volume or area to produce hydrocarbons.

Additionally, algae species may either have growth-associated or growth-disassociated hydrocarbon production, in which the growth-associated hydrocarbon production increases proportionately with cell growth and growth-disassociated production has hydrocarbon production that is not dependent on cell-growth. *B. braunii* is a growth-associated hydrocarbon producing species while *C. vulgaris* is growth-disassociated, as shown in **Figure 6**, in which a majority of its lipids are produced after cell growth stagnates.

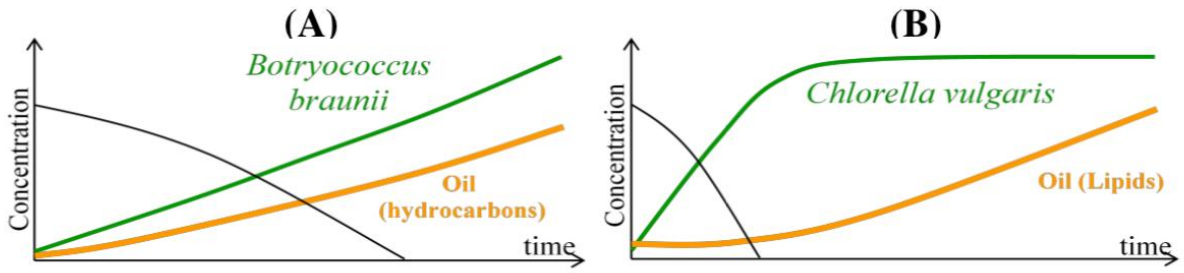


Figure 6. Growth-associated (a) and growth-disassociated (b) hydrocarbon production for B. braunii and C. vulgaris adapted from Tuerk (2011).

Specific productivity removes the unfair advantage of higher cell count on productivity by measuring the rate of hydrocarbon production based on the algae biomass, which results in a productivity that is more intrinsic to an algae species. As described within **Table 1**, there are multiple methods to determine the amount of biomass within the reactor, but since biomass dry

weight is most widely presented in data and one of the more accurate estimations of reactor biomass concentration in comparison to OD and chlorophyll measurements, this thesis will calculate specific productivity in terms of biomass dry weight.

For hydrocarbon concentrations in terms of grams hydrocarbon per reactor volume or area, dividing by the corresponding biomass concentration in the reactor during the same sampling time will yield hydrocarbon production mass per biomass dry weight since the equivalent reactor volume or area will divide out. If hydrocarbon concentrations are based on solvent volume post-extraction, then determine the accumulation of hydrocarbon mass using the solvent volume used during extraction and then divide by biomass at the corresponding time of hydrocarbon extraction within the reactor using the reactor volume or area. Note that since changes in hydrocarbon will reflect similar efficiencies of extraction, one can obtain a productivity that is reasonably accurate, even if there is incomplete extraction.

For continuous systems, determine the volumetric or areal hydrocarbon productivity using **Equation 8** and then divide by the steady operational biomass concentration to yield a specific hydrocarbon productivity that is fundamentally more accurate than attempting to take derivatives associated with time-variant batch or fed-batch operation.

Conversion from Total Algae Growth Period to Photohours and Incident Photons

As described within the Chapter 3, the algae culture growth curves are often plotted over the entire culture growth period. To normalize the light/dark growth cycles the algae are grown under, the time aspect of the biomass growth curves can be converted to photohours, which are the total number of hours the algae culture receives lighting.

For algae grown under controlled, ‘indoor’ conditions, where lights are switched on and off in the matter of seconds, the growth period is multiplied by the number of hours of light the algae receive per day. The conversion between growth period in days to photohours for different light/dark cycles is shown in **Table 2**.

Table 2. Conversion between Growth Period and Photohours for Various Light/Dark Cycles

Growth Period (days)	Continuous lighting - 24 hours light (photohours)	16/8 hour light/dark cycle (photohours)	12/12 hour light/dark cycle (photohours)
0	0	0	0
1	24	16	12
2	48	32	24
3	72	48	36
4	96	64	48
5	120	80	60
n	24n	16n	12n

For open-culture systems under inconsistent lighting, such as sunlight, the amount of light the algae culture receives depends on the time of day and the angle of the sun, which means photohour requires integration of incident photons on the system with time and then a conversion from photons to photohours using reactor surface area. Instead, in order to compare light levels for systems that have constant or inconsistent lighting, cumulative incident photons available to the algae can be calculated by integrating either PPFD or the photosynthetic photon flux (PPF) of the light source, which is the rate of photon produced by a light source within the PAR range in

units of µmol/s, with respect to time. PPF of a light source is not often indicated within literature or listed on light products, so PPFD can be used instead.

For algae grown under a constant light source with defined light/dark cycles, the cumulative incident photons can be calculated by multiplying the light intensity (PPFD) by the lighted reactor surface area (A) and the time of light availability (photohours), as shown in **Equation 13**.

$$Cumulative\ Incident\ Photons = PPFD * A * Photohours \qquad \textbf{Equation 13}$$

For inconsistent light sources like sunlight, the PPFD from the sun can be logged over time using a PAR sensor normal to the Earth to yield a graph of light intensity versus time of day, as shown in **Figure 7**.

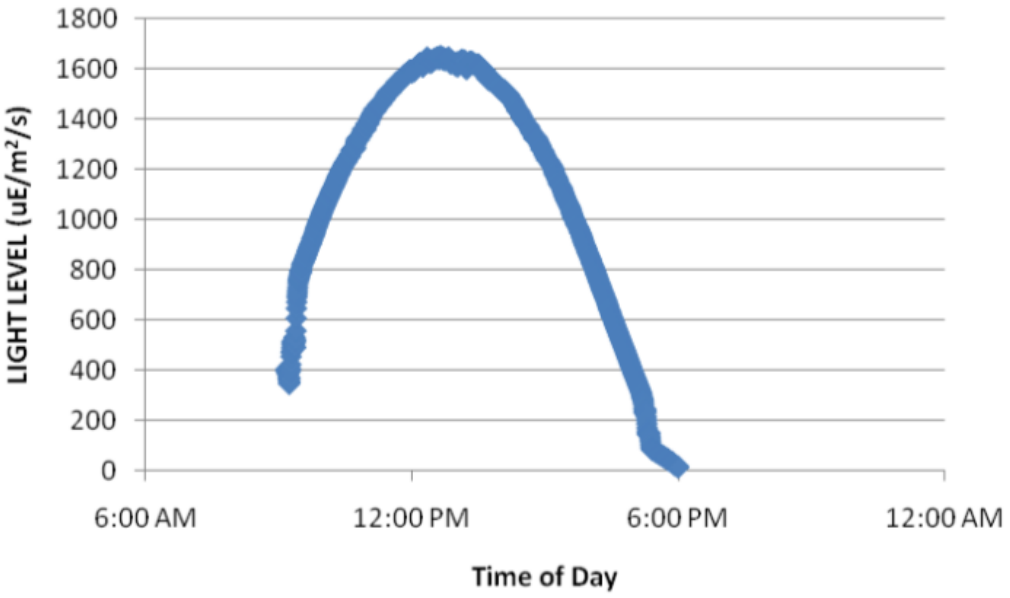


Figure 7. Adapted from the Photosynthetically Active Radiation (PAR light) levels logged at the top of the outdoor trickle-screen photobioreactor (Sunny day, early November) from NSF-SBIR Research for award #0945592[14].

Since the light intensity measurements are taken normal to the Earth from the top of the reactor, the position of the sun over time must be determined in order to calculate the incident

photons on the reactor surfaces. Using positional astronomy equations, the solar zenith, elevation, and azimuth angles can be calculated as shown in **Figure 8**. These calculations were applied to a horizontal thin trickle-film bioreactor so that flux can be applied based on the angle if incidence on that surface based on the position of sun in the sky.

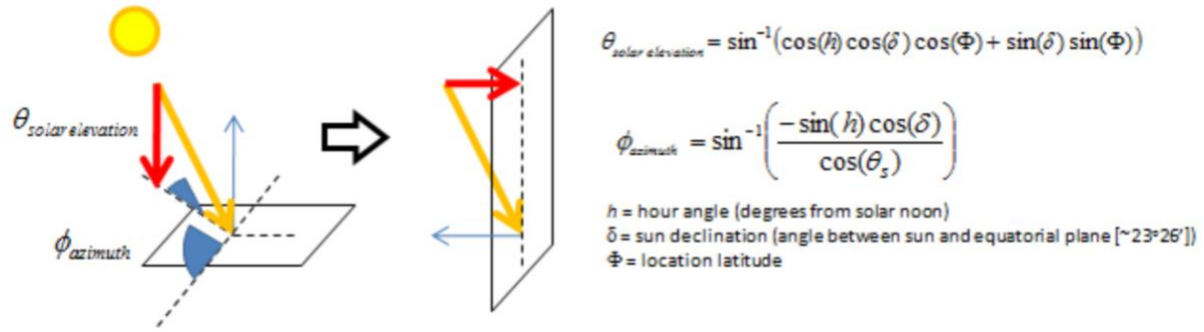


Figure 8. Adapted from NSF-SBIR Research for award #0945592 depicting the procedure developed by Brandon Curtis to calculate incident photons in a given day for a vertical trickle-screen photobioreactor where light intensity measurements are taken normal to the Earth[14].

Based on the angle of solar elevation, the atmospheric refraction effects that alter the apparent position of the sun in the sky can be corrected. Under the assumption that light intensity can be approximated as a direct solar beam, combining direct sunlight and scattered light into one vector, the corrected solar elevation and azimuth angles can then be used with trigonometric properties to calculate the total light intensity over time on a given reactor surface, as shown by the yellow arrow in **Figure 8**. The corrected light intensity normal to the Earth from Figure 2 as well as calculated light intensities over time for the North-South and East-West Reactor Surfaces are shown in **Figure 9**.

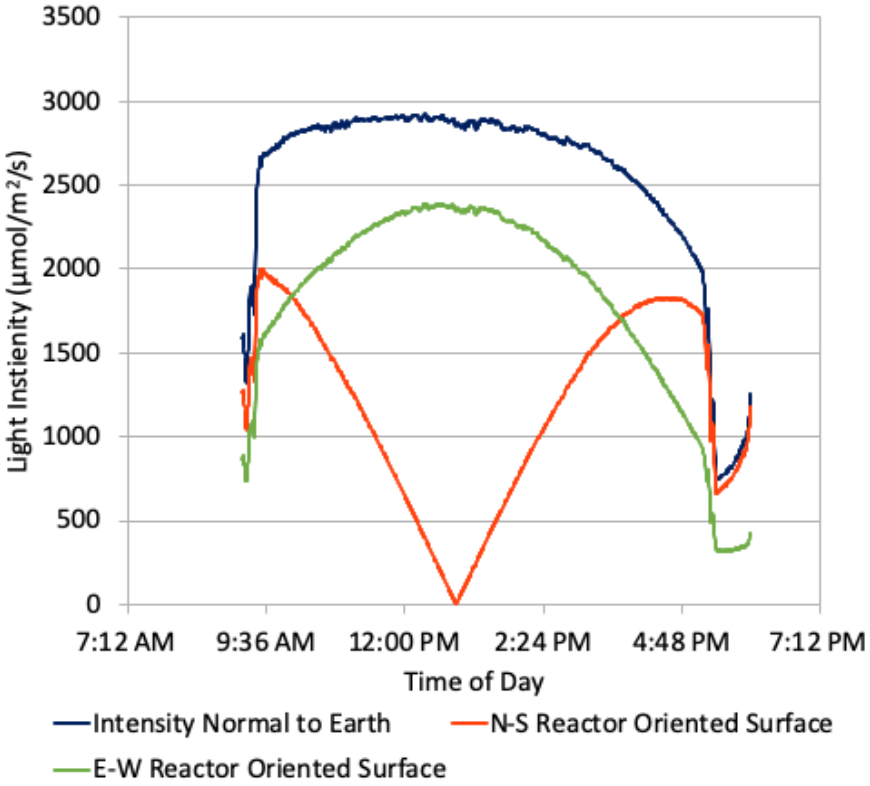


Figure 9. Corrections for sunlight intensity measurements normal to the Earth as well as on the North-South and East-West oriented reactor planes. Adapted from NSF-SBIR Research for award #0945592[14]

An integral of the light intensity with time for the different surfaces in **Figure 9** yields the micromoles of incident photons per area. Multiplying each areal incident photon value by the total lighted reactor surface for each side of the reactor yields the cumulative incident photons for each side of the reactor, when summed together results in a cumulative incident photon value in micromoles that can be compared to the cumulative photons determined for constant lighted systems using **Equation 13**.

Calculation of Photon Energy Input

Notably, the biologically relevant method of determining photon energy input is to measure the energy from photons actually absorbed, known as the biologically accessible

photons (BAP), by the algae rather than the incident PPFD. The method for calculating BAP is shown within the thesis by Nugent (2020), in which Beer-Lambert Law is used to determine a correction factor to measure how many incident photons are transmitted through the reactor and can be used for photosynthesis.[8] The relative light intensity emission spectrum for the bulbs used to illuminate the algae culture is then plotted over the PAR wavelengths to determine photon flux at each wavelength. The emission spectrum is then overlaid on the algae absorbance spectrum at a robust wavelength that maximizes light scattering and a sensitive wavelength that is dominated by light absorbance to determine the range of PAR absorbance from the total incident photons. An average of the two PAR absorbance values can be used within the Beer-Lambert law to determine BAP.

Few literature articles report BAP so the photosynthetic efficiencies in this thesis are calculated using the PPFD of the light source. As mentioned within the Productivity Definitions chapter, PPFD must be converted to photosynthetically active irradiance to determine the energy available from the incident PPFD. An average conversion factor for light energy associated with PAR wavelengths underwater, is 0.24 MJ per mole quanta.[6] The underwater light energy value can thereby be used to convert the total moles of incident photons from **Equation 13** into total photon input energy. **Equation 11** can then be used to calculate photosynthetic efficiency for the algae based on photosynthetically active irradiance.

The method of finding total incident photons for systems under time-variant lighting is described in the previous section. Photon energy can be determined by wavelength:

$$E_{photon} = \frac{hc}{\lambda} \qquad \textbf{Equation 14}$$

where h is Planck's constant, c is the speed of light, and λ is the wavelength of light, usually the robust wavelength for the algae species. **Equation 14** yields photon energy in terms of kilojoules

per mole, so multiplying by the total incident photons in a day yields total photon energy into the reactor system in a day.

Chapter 4

Algae Growth Data Digitization

A major component of this thesis is dependent on software, e.g. Engauge Digitizer, developed by Mark Mitchell et al[7], to obtain values from graphed hydrocarbon concentrations and *B. braunii* growth time curves from literature. As a result, the biomass and hydrocarbon concentrations are dependent on the accuracy of how the software or the user selects the data point as well as the functions programmed into software that connects the points. Engauge Digitizer allows the user to define the axes on an imported chart and then manually place points onto the curve and the program will create x and y coordinates for those points based on the defined axes. This method of point generation is suitable for simple productivity calculations, especially to check specific biomass concentration values given within the literature or on the graph itself as well as determining average productivity. Calculating instantaneous productivity, on the other hand, requires a large number of points on a given curve to minimize delta x and obtain the derivative at a given point, as defined in the second portion of this chapter.

Tabularizing Graphical Data from Literature Data

Once downloaded via one of the options described in the Engauge Digitizer GitHub page, the application will open to a blank screen as shown in **Figure 10**.

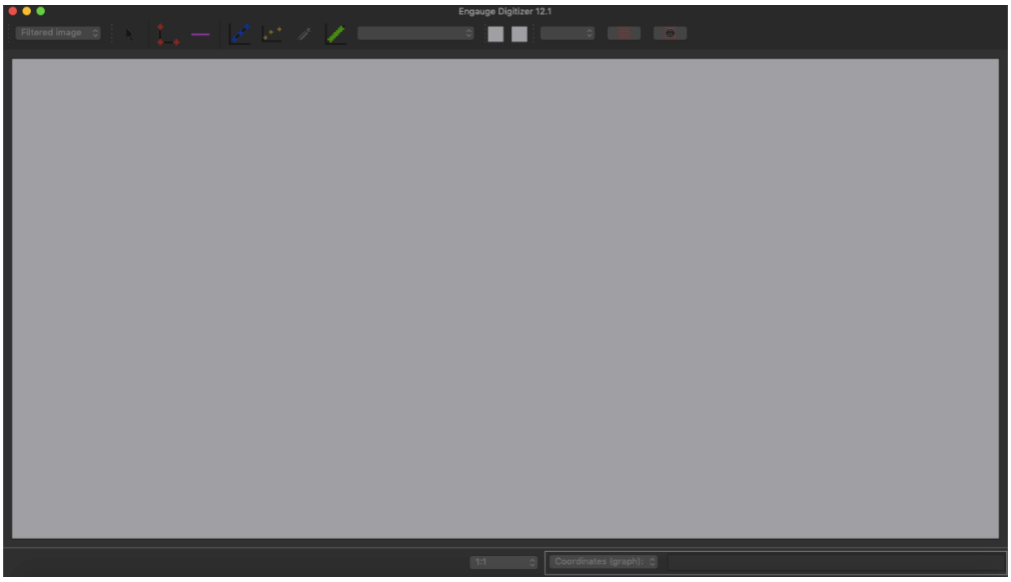


Figure 10. Landing screen for Engauge Digitizer Software

An image of any graph to be tabulated can be imported into the software by going to the File tab and using the “Import…” option for a single graph with two axes. For more complicated images, such as multiple coordinate systems in one image or graphs that have two coordinates defined on each axis, use the “Import (Advanced)...” option.

Within this example, the hydrocarbon accumulation and biomass growth time-curve graph from Casadevall et al (1985) will be digitized, shown in **Figure 11**.

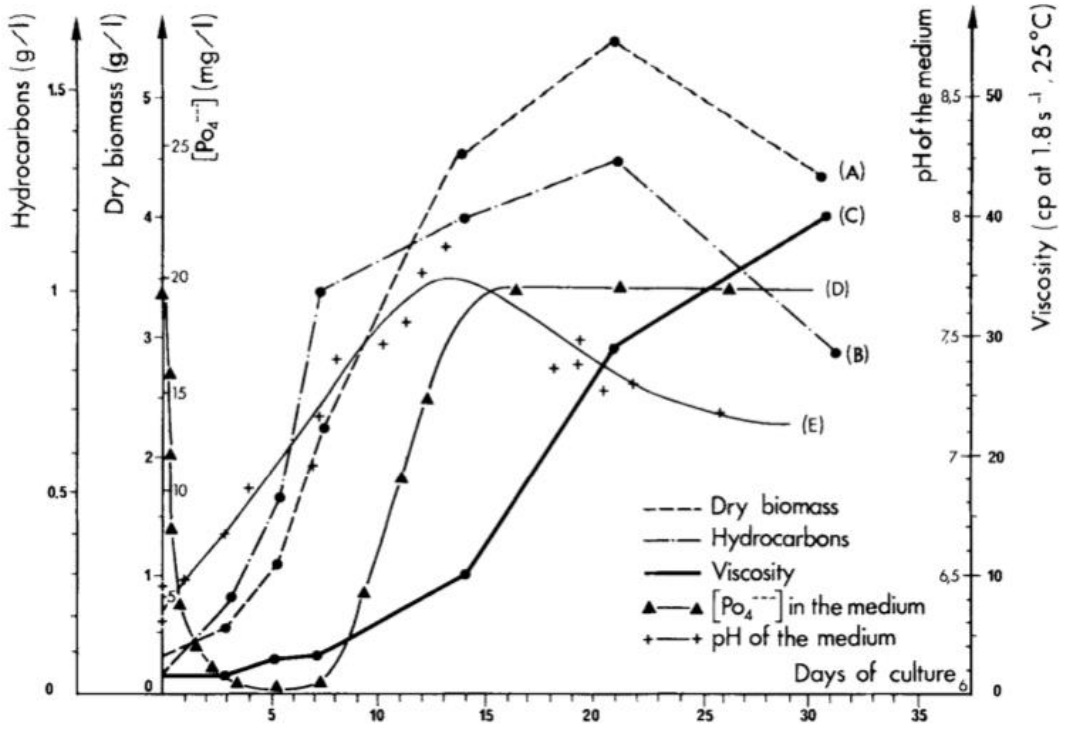


Figure 11. Dry biomass and hydrocarbon accumulation time curves from Casadevall et al (1985) air-lift batch cultures[17].

The curves of interest are on two separate coordinate systems, so the advanced import was used for this graph. Once the image is imported, a Checklist Guide Wizard appears, requesting that the user names the curves and choose whether the curves are drawn with lines or only points, shown in **Figure 12**. Since the biomass growth and hydrocarbon accumulation time curves have lines connecting the data points, the "With lines (with or without points)" is chosen for both coordinate systems.

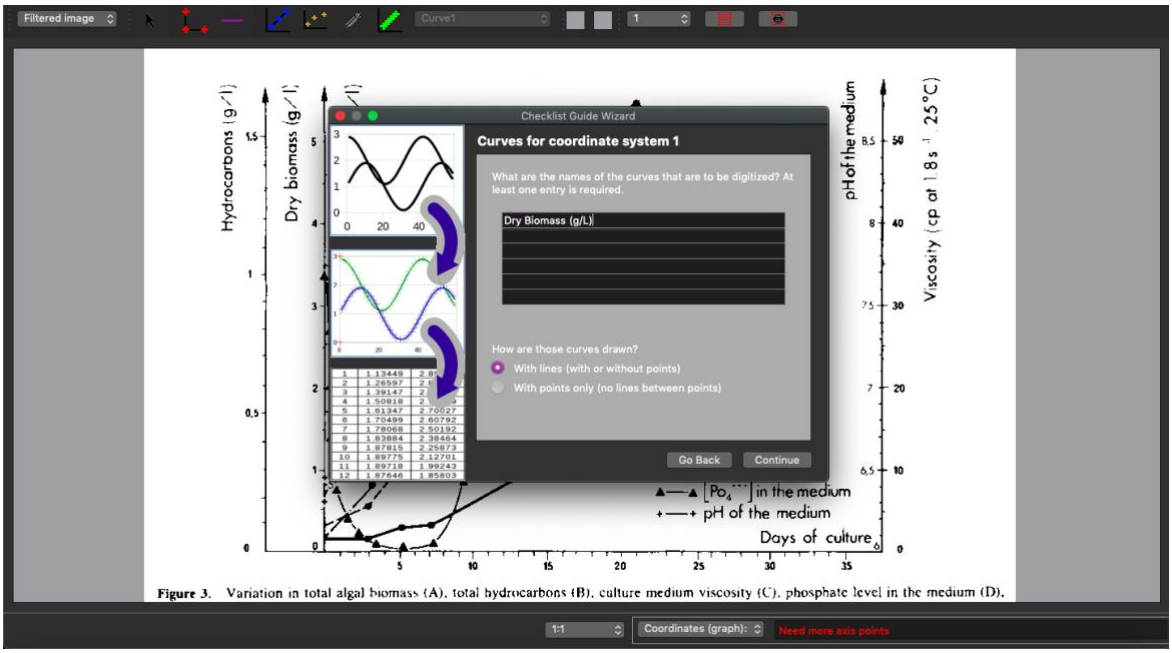


Figure 12. Checklist Guide Wizard within Engauge Digitizer with Dry Biomass curve named for tabulation for coordinate system 1 (top) and hydrocarbon accumulation for coordinate system 2 (bottom).

After completing the Checklist Guide Wizard, the graph is shown with a Checklist Guide on the right-hand side shown in **Figure 13.** If the Guide does not appear, it can be enabled through the View tab and "Checklist Guide Toolbar" option.

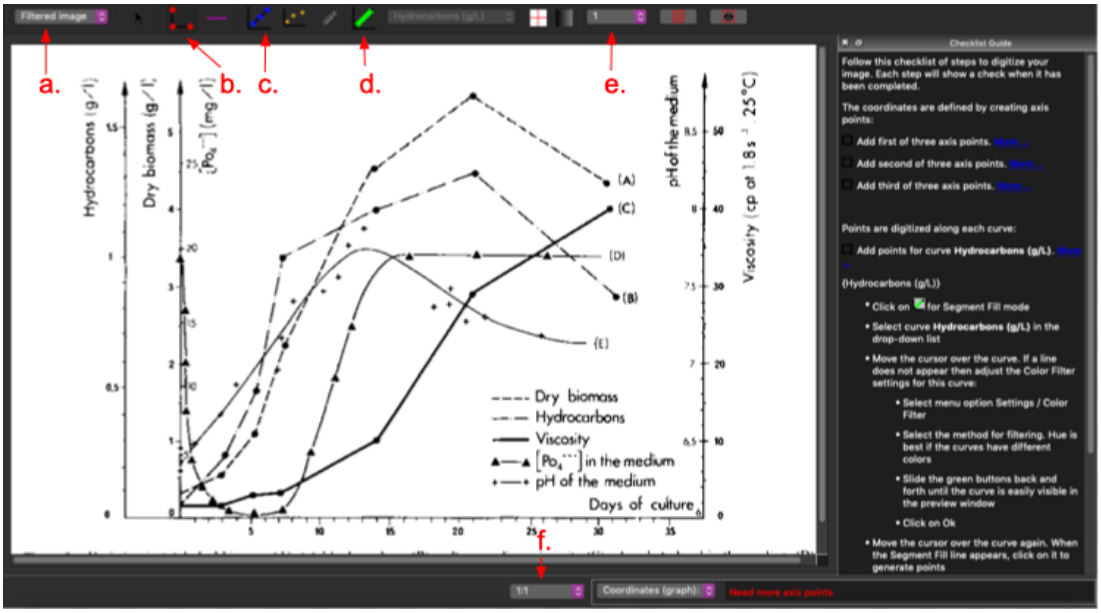


Figure 13. View of Filtered Graph and Checklist Guide in Engauge Digitizer once the graph is imported into the application.

In **Figure 13**, letter a is the filtered image that Engauge Digitizer generates based on the image that is imported into the software based on the color and resolution setting in the software. If lines are not visible in the "Filtered Image" option, then Engauge Digitizer's filter can be adjusted using the "Color Filter" option under the Settings Tab in the program. Letter b is the tab to define the axes for the graph, letter c is to select points on a given curve, letter d is the Segment Fill tool that allows users to generate points along a curve that are evenly spaced, letter e allows the user to select which coordinate system they are defining, and letter f allows the user to zoom into the graph.

As indicated by the Checklist Guide in **Figure 13**, the first step is to define the axes by clicking the button indicated by letter b in **Figure 13**. This will generate red crosses on the graph. The user should define the axis by clicking on three areas on the axis and inputting the values when prompted. Typically, the intersection point between the axes and the maximum x and y values on the axes should be defined, as shown in **Figure 14**. Note, while the hydrocarbon concentration has a different scale than the dry biomass concentration, the y-axis for the dry biomass should be used to digitize the hydrocarbon concentration values since day zero for the experiment intersects with 0 g/L of the dry biomass y-axis.

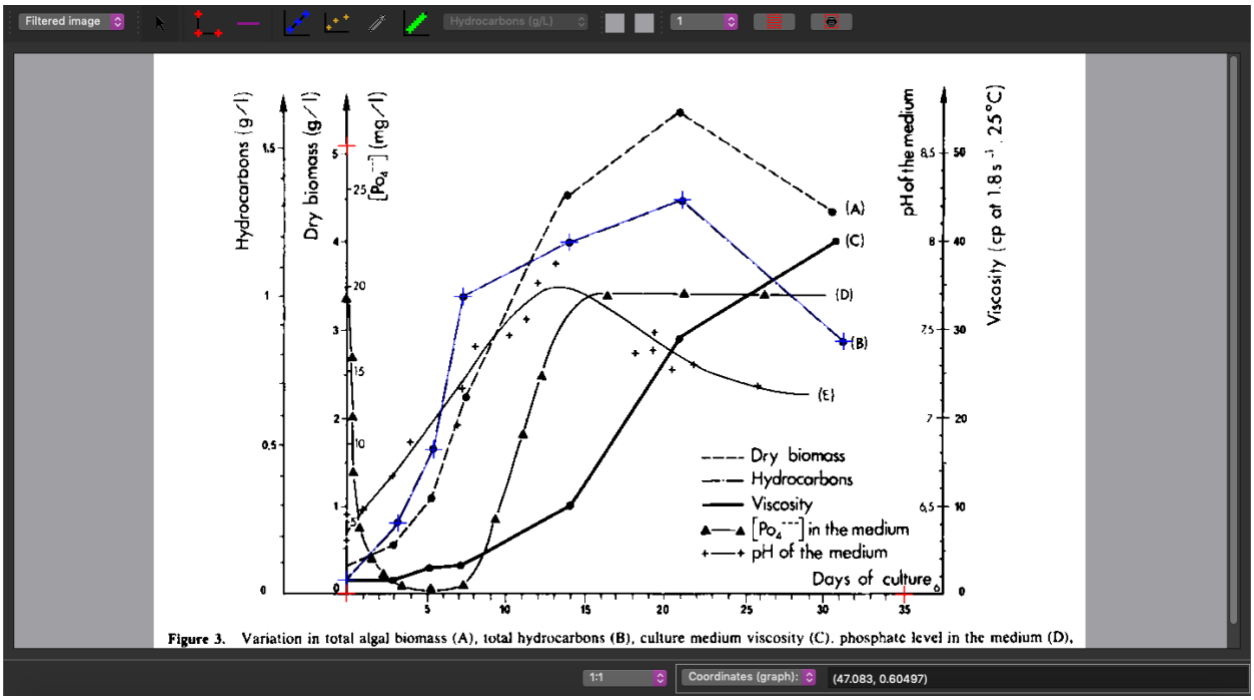


Figure 14. Defined axis points and digitized points along the hydrocarbons curve for the algae growth curve from Casadevall (1985).

Figure 14 also has the points along the hydrocarbons curve defined via the same process as the axes (shown with the blue crosses), but the values do not have to be explicitly defined. Notably, the curve points can be altered within the "Curve Properties" under the Settings Tab and the line that connects the point can be switched between straight or smooth functions or relations, as shown in **Figure 15**. The relation option allows the user to generate points anywhere on the graph in any order, but the function option restricts the curve to have only one y value for any given x value. The straight option generates straight lines between the digitized points while the smooth option connects the points with a curve. Within this example, the "function - straight option" was selected since the hydrocarbon curve has only one y value for any given x value and also has straight lines connecting the point within the original graph.

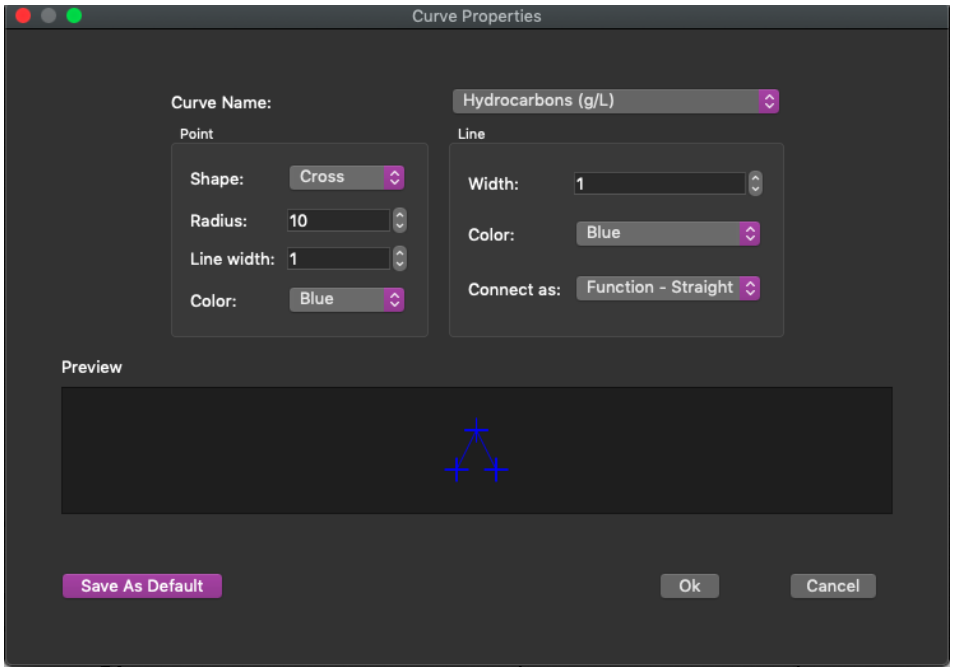


Figure 15. Curve properties for digitized curve points within Engauge Digitizer.

With the axes defined and hydrocarbon points digitized, the digitized points can be exported into a CSV file using the "Export…" option within the File tab. The file can then be imported into Excel or another program for productivity calculations.

The digitization process was repeated for the dry biomass curve. The digitized dry biomass graph is shown in **Figure 16**.

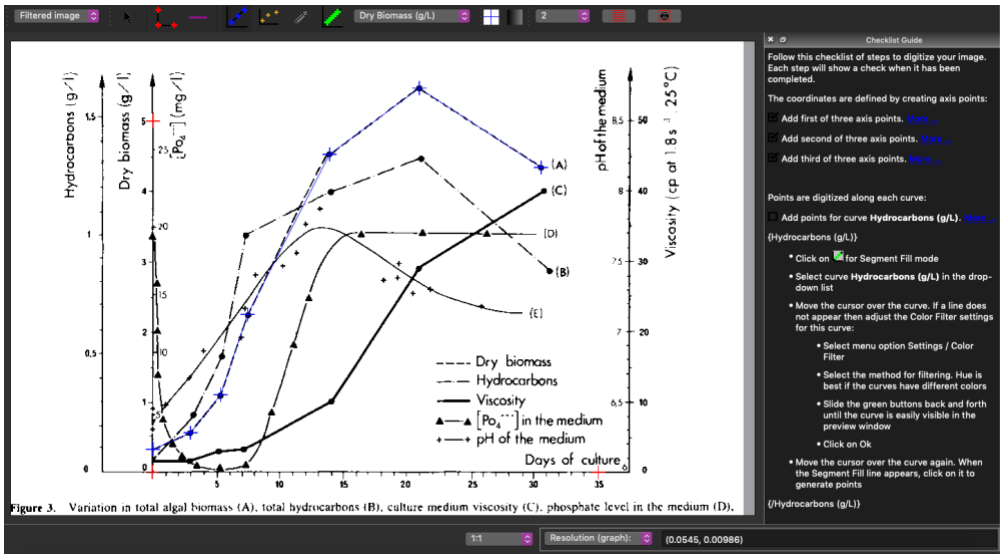
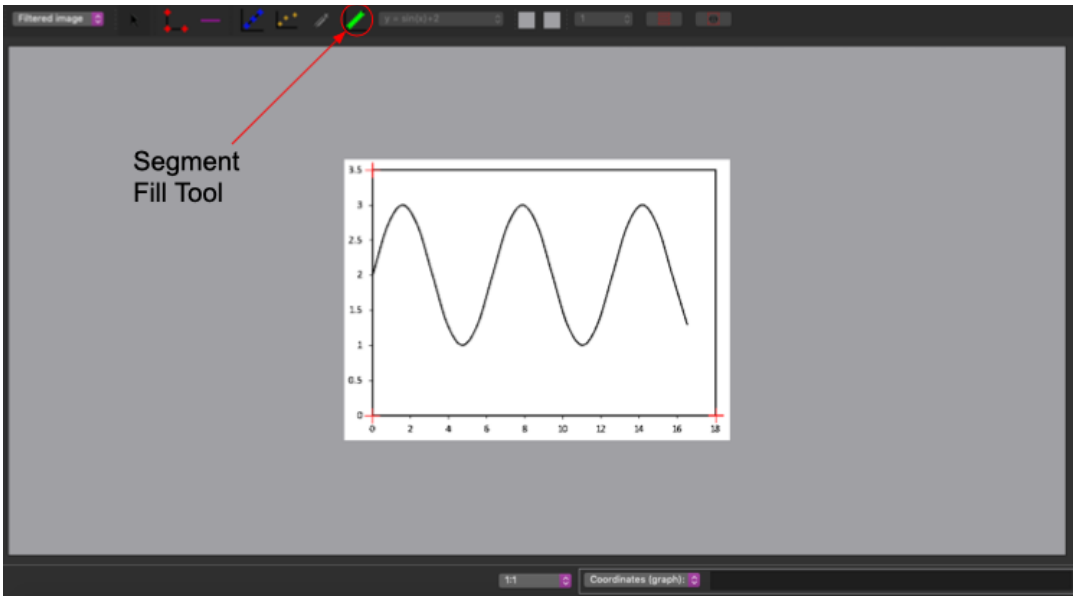


Figure 16. Axes definition and dry biomass digitized points for Casadevall (1985) algae growth curve.

Generating Points along Algae Biomass Curves for Instantaneous Productivity Calculations

As described within Chapter 3, the fitting of an equation to an algae growth time curve brings inherent constraints, especially when a derivative is taken. Therefore, instead of fitting an equation to a growth curve, Engage Digitizer can be used to generate any number of points along a given growth curve as long as the curve can be detected by the Digitizer's scanner. Therefore, both digital and hand-drawn curves can be digitized. For graphs that have linear connections between the points, such as the Casadevall (1985) hydrocarbon accumulation and biomass production curves in **Figure 11**, a curve should be hand-drawn and then digitized. Visually fitting a curve to the data points minimizes error in both the x and y directions and when replicated by another person, the curve will usually be drawn similarly for the same data points. Therefore, to calculate the instantaneous productivity for curves connected with straight lines, smooth curves will be hand-drawn and digitized within this thesis.

Within Engauge Digitizer the "Segment Fill" tool can be used to generate points over a graph which can then be tabulated. The software recognizes the entire curve using the Segment Fill tool and can generate points along the curve based on the user-specified pixel separation and a few other parameters as shown in **Figure 17**. Altering the line width within the Segment Fill settings or the point radius or line width within the Curve Property settings do not alter coordinates of the points generated along the curve, so the" Point Separation" and "Fill Corner" settings under the Segment Fill settings are the main adjustable parameters when calculating digitized derivatives.



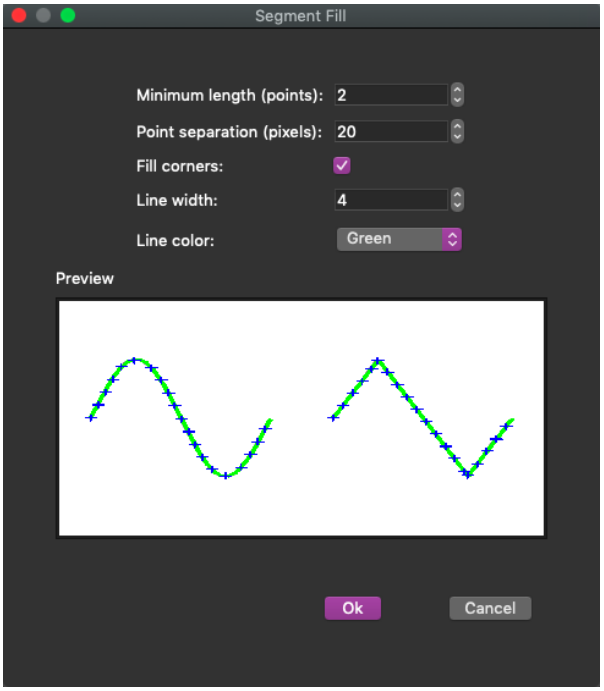


Figure 17. (Top) Segment Fill tool in the Engauge Digitizer main window that allows users to generate points along a curve. (Bottom) The Segment Fill settings in Engauge Digitizer allow users to specify the minimum length between points, the pixel separation, whether the points should fill corners on curve, and line parameters.

If the curve crosses one of the axes (entering a different quadrant), then the Segment Fill will recognize the curve as two or more pieces on either side of the axis, as shown in **Figure 18**.

For growth time curves, this should not be an issue since concentration and time should usually be on the positive x and y axes.

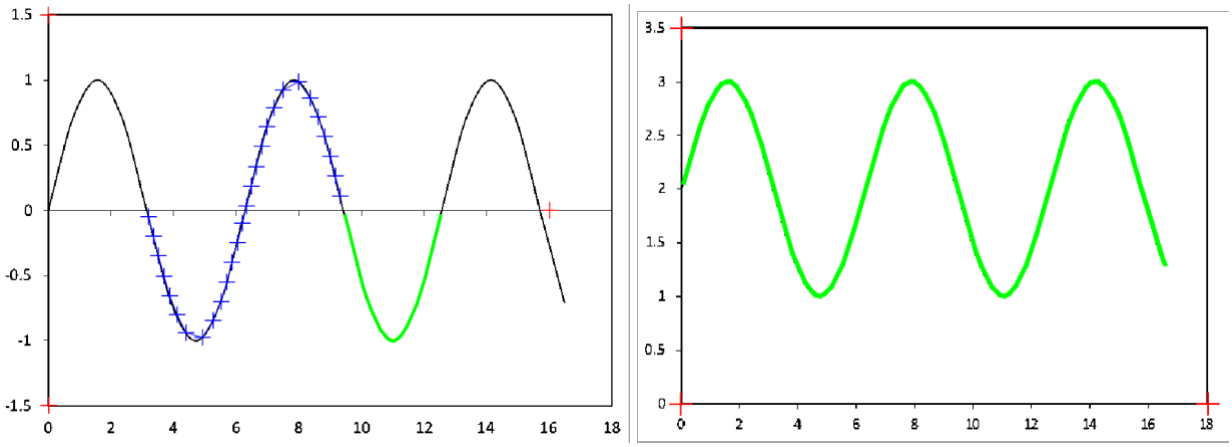


Figure 18. Digitization of Excel generated graph of y = sinx. The x-axis splits the curve into segments as shown by the section highlighted green using the segment fill tool (left). For curves completely within the first quadrant, the entire graph is recognized and highlighted (right).

Due to the adjustable parameters for point separation as well as uncertainty in the accuracy of the point generation along the curve, an experiment was done by digitizing a graph with a known function to determine how varying the point separation within Engauge Digitizer affects the calculated derivative from the digitized data.

To test the Segment Fill settings, a sinusoidal curve was used within the experiment since it has a simple derivative to compare to the derivatives generated from the digitized curve and also shows whether the point generation is consistent over the entirety of the graph since the function oscillates. To avoid the curve crossing the x-axis and fragmenting under the Segment Fill tool, the sinusoidal graph was shifted up two units to yield a final equation of y = sin(x) + 2, which has a derivative of cos(x).

Using the digitized excel plot of y = sin(x)+2, the “Point Separation (pixel)” setting under the Segment Fill section of Engauge Digitizer’s settings was altered. The total desired number of points to be generated along the graph could not be specified for without trial and error with the

Point Separation tool so Point Separation settings of 20, 15, 10, and 5 were chosen which yielded a total number of 58, 75, 112, and 223 points respectively along the curve (N). Altering the point separation settings could generate a specific number of points (e.g. N = 100, 50, etc) if desired, but the range of points given by altering the Point Separation setting by intervals of 5 yielded a sufficient range of generated number of points to represent the impact of changing the number of digitized points.

The digitized points were imported into Excel, where the x and y values were plotted. The equation $y=\sin(x)+2$ was used to generate the actual y values for the generated x values. Both the digitized and calculated y values were plotted against the average x value for all values of N. Though not perfect due to human error in the axis definition as well as the Segment Fill Tool having slight variations in where the points are generated along the curve, the generated points provided a good match to the calculated values as would be expected for a digitization procedure.

The derivative was calculated using the generated coordinate points via a simple $dy/dx=(y_i-y_{i-1})/(x_i-x_{i-1})$ calculation which was plotted against the averages of the x_i and x_{i-1} and values. The actual derivative was calculated using $y'=\cos x$ using the x_i values and also plotted within the same graph, shown in **Figure 19** for N=58, 75, 112, and 223.

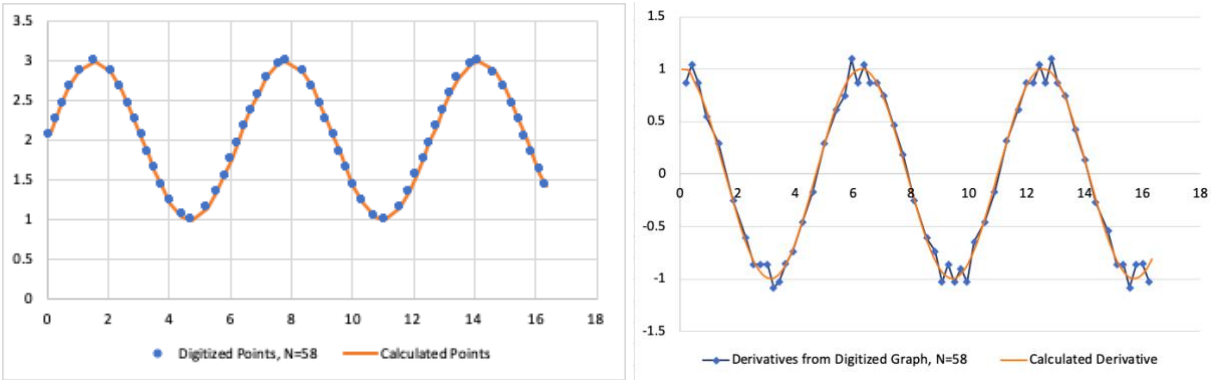
Results and Discussion

Digitized Graph

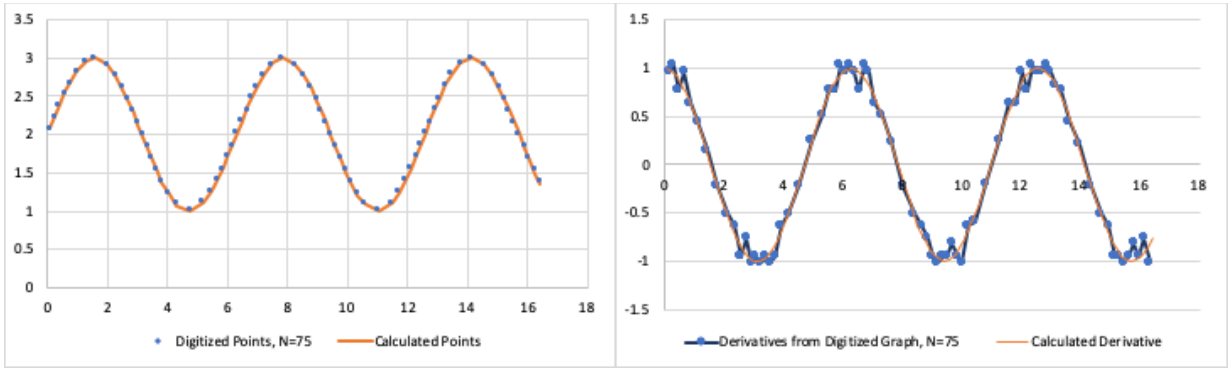
Regardless of the N value, all the digitized graphs had the same difference between the actual y value and the digitized value for any specific x value, meaning that altering the Point Separation setting does not alter the accuracy with which the points are generated along the graph.

Calculated Derivatives from Digitized Data

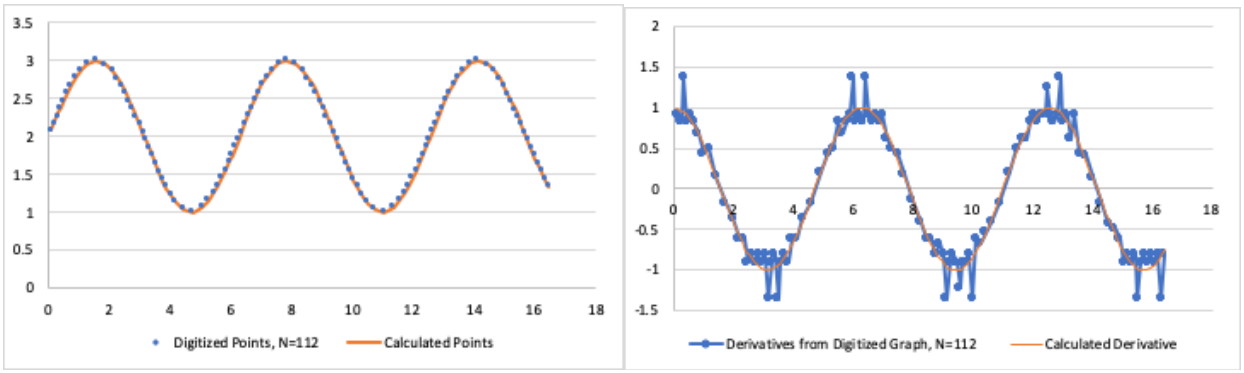
a.



b.



c.



d.

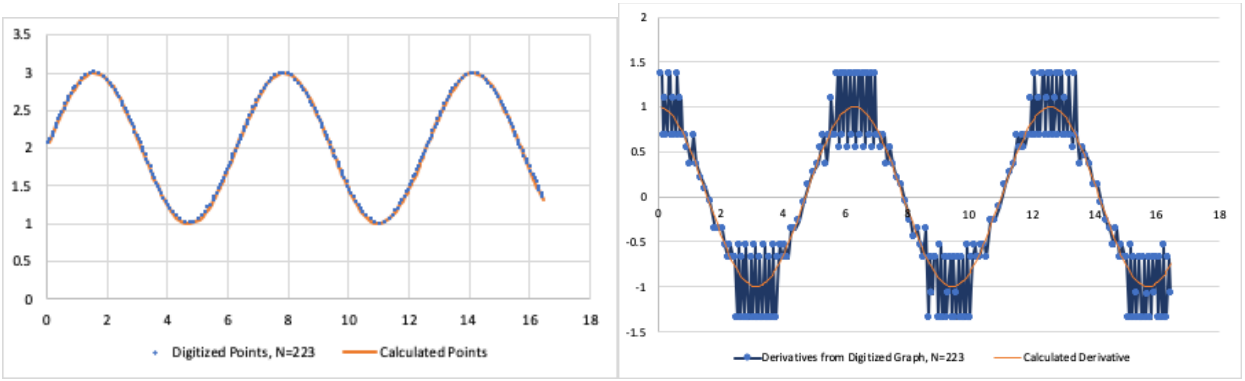


Figure 19. Digitized points on sinusoidal curve y=sin(x)+2 via Engauge Digitizer on left and the respective derivatives from the digitized data points on the right. Figures A through D correspond to N=58, 75, 112, and 223, which are the number of points generated along the curve.

The digitized data points follow the actual curve well for all the N values as shown in the graphs on the right in **Figure 19**. As described in the methods section, the derivatives were determined by calculating the change in y divided by change in x between two points, as shown in **Equation 15**, where the fundamental definition of a derivative would dictate minimizing dx to obtain a value close to the slope of the line tangent to the curve.

$$Digitized\ Derivative = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \qquad \textbf{Equation 15}$$

However, when working with digitized data rather than a smooth function, minimizing dx was found to inherently amplify any inaccuracies in dy, so when taking the slope of two points very close together, the change in y values will magnify. Within **Figure 19**, the left-hand graphs show this issue at the maximum and minimum derivative values, which correspond to inflection points in the digitized data with smaller change in x between the data points. The magnitude of the fluctuations of the derivatives also increases as N increases, in which the graphs of N=122 and 223 have derivatives nearing 1.5 (or 50% error) when the actual derivative value is one. Higher N values mean less separation between the generated data points, which becomes the basis of the fluctuations. Since the generated data points are closer, dx is smaller and small variations in dy that result from digitization are amplified into erroneous calculated derivatives. To provide for a quantitative index of how close the digitized derivative curves are to the actual values, a

sum square difference was taken and then normalized by dividing by the total number of data points as shown in **Equations 16** and **17**.

$$Total\ sum\ square\ difference = \sum_{i=1}\left(\frac{dy}{dx_i} - \frac{dy}{dx_{actual}}\right)^2 \quad \textbf{Equation 16}$$

$$Normalized\ sum\ square\ difference = \frac{\sum_{i=1}\left(\frac{dy}{dx_i} - \frac{dy}{dx_{actual}}\right)^2}{N} \quad \textbf{Equation 17}$$

The total and normalized sum square differences for the sinusoidal curves shown in Figure 3 are shown in **Table 3,** in which both the total and normalized sum square difference increase with the amount of fluctuation in the digitized derivative curves.

Table 3. Overview of the digitized sinusoidal graph y=sin(x)+2 with varying number of points generated along the graph as well as whether the generated y values are not statistically significantly different from the actual y values.

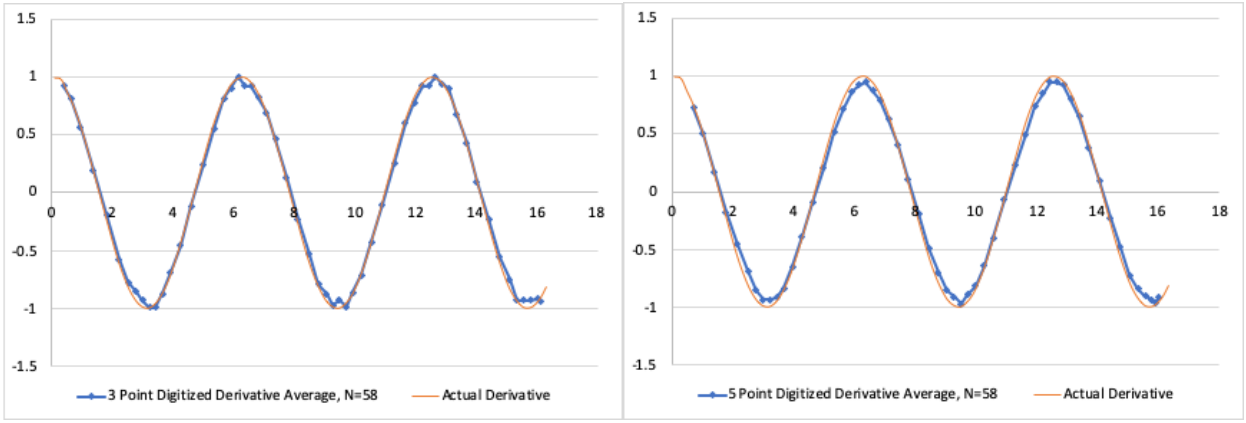
Point Separation (Pixels)	Number of Points Generated Along the Graph (N)	Total Sum Square Difference between Digitized Derivative Points and Actual Derivative	Normalized Sum Square Difference between Digitized and Actual Derivative Values
20	58	0.0342	0.00601
15	75	0.529	0.00715
10	112	0.772	0.696
5	223	19.2	0.0863

In order to minimize the fluctuations, a three-point and five-point average of the derivatives can be taken to yield derivative values closer to the actual derivative. For the three-point average, the derivative for a given x value is averaged with the subsequent and previous derivative values, shown in **Equation 18**.

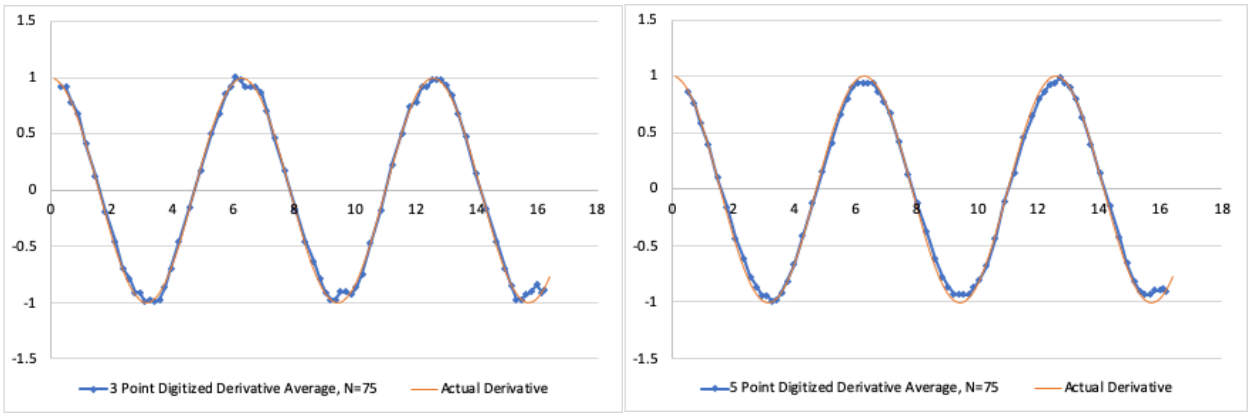
$$Three\ point\ average\ derivative = \frac{\frac{dy}{dx_{i-1}} + \frac{dy}{dx_i} + \frac{dy}{dx_{i+1}}}{3} \quad \textbf{Equation 18}$$

For the five-point average, two points after and before the selected derivative value are averaged with the selected value. The averaged digitized derivative points are graphed against the averaged x values of the averaged digitized derivative values. The results of the three and five-point averages are graphed in **Figure 20** for the four N values selected in **Table 3** and **Figure 19.**

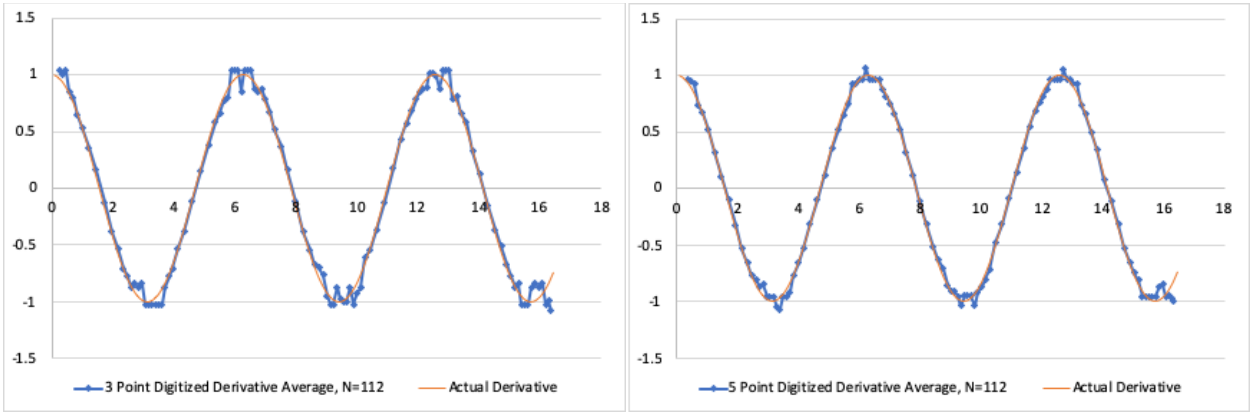
A.



B.



C.



D.

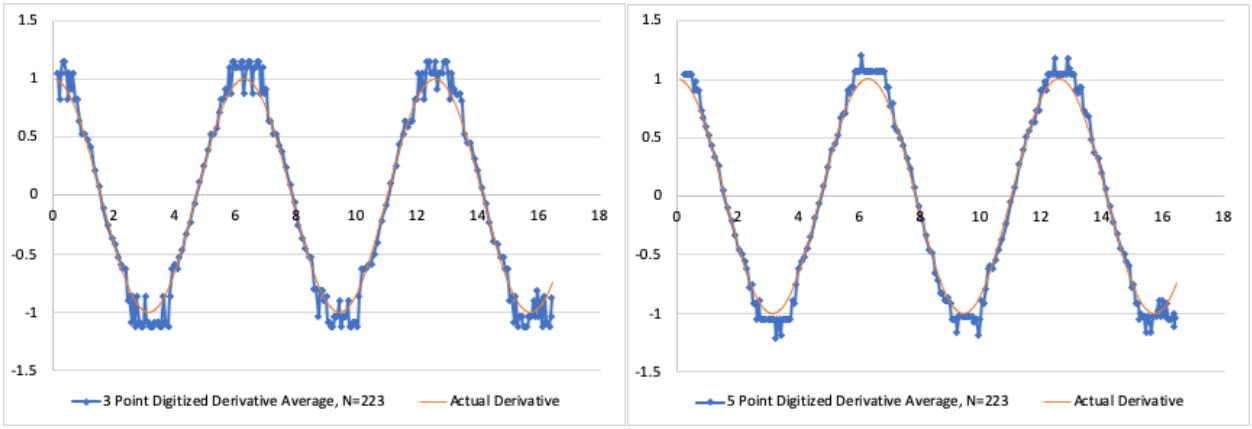


Figure 20. Three-point and five-point derivative averages for each point separation value A to D described in **Table 3** for the sinusoidal curve.

Graphically, when compared to the original digitized derivative curves in **Figure 19**, all of the digitized derivative curves in **Figure 20** more closely align to the actual derivative curve, which means that the process of averaging a given point with its neighbors effectively minimizes the fluctuations seen in **Figure 19.** However, for graphs with fewer points generated along the curve, the digitized derivative values in the five-point average curve appear further from the actual derivative curve than the three-point digitized derivative averages. To confirm this numerically, a sum square difference between the averaged derivative values and the actual derivative was performed to find how the three and five-point averages compared for the four N values. Of note, the actual derivative was calculated from the three or five-point averaged x

values to eliminate digitized derivative offset in the x direction. This means that the actual derivative value from **Equations 16** and **17** when calculating the sum square difference was calculated using the three or five-point average x value that the averaged digitized derivative value was plotted against. The sum square difference values for all the N values as well as normalized sum square difference values, in which the original sum square difference is divided by the number of average digitized derivative points, are shown in **Table 4**. Note for the normalized sum square difference for digitized derivative averages, the total sum square difference is divided by the number of averaged points, not the total number of points generated along the curve in Engauge Digitizer (N).

Table 4. Total and Normalized Sum square difference for the three and five-point average digitized derivatives for the graph y=sin(x)+2

Point Separation (Pixels)	Number of Points Generated Along the Graph (N)	Total Sum Square Difference between Digitized and Actual Derivative Values (3 point average)	Normalized Sum Square Difference between Digitized and Actual Derivative Values (3 pt avg)	Total Sum Square Difference between Digitized and Actual Derivative Values (5 pt avg)	Normalized Sum Square Difference between Digitized and Actual Derivative Values (5 pt avg)
20	58	0.0727	0.00130	0.174	0.00317
15	75	0.122	0.00167	0.116	0.00161
10	112	0.507	0.0046	0.221	0.00203
5	223	3.30	0.0150	2.21	0.0101

Both the total and normalized sum square difference for the three-point average for N=58 is lower than the values for the five-point average, meaning that the digitized values for the three-point average more closely align to the actual derivative values. At N =75 points, the total and normalized sum square difference are similar, though the five-point average has a slightly

lower sum. At N = 112 and 223 points, though, the sum square difference for the five-point averages is far lower than the three-point averages. Of note, the lowest sum square difference value for the sinusoidal curve occurs with the lowest N value and three-point average.

The variation in accuracy for three-point and five-point averages lies in the separation between the points at the different N values. At lower N values, the points are further separated and averages multiple neighboring points will shift the points towards the neighbors, especially at the turning points in the derivative graph since the neighboring points on both sides of a given point are either above or below the given point. Therefore, at lower N values, a three-point average should be used and at higher N values, five-point averages are more accurate since more points are averaged together to bring the fluctuating points closer to the actual derivative and the cluster of points are in the same small region of the curve, while the points for a five-point average at low values of N correspond to a larger portion of the curve so taking the average of the derivative will more likely shift the value towards the adjacent points.

These observations suggest that for higher N values, increasing the number of points averaged with a given point should bring the digitized derivative closer to the actual value. To test the theory, a seven-point average was conducted with the sinusoidal graph of N = 223. This yielded total and normalized sum square differences of 1.84 and 0.00841, which is lower than both the three-point and five-point average values in **Table 4.** The fluctuations are also minimized the most with the seven-point average, as shown in the graph of the seven-point average for N=223 in **Figure 21**.

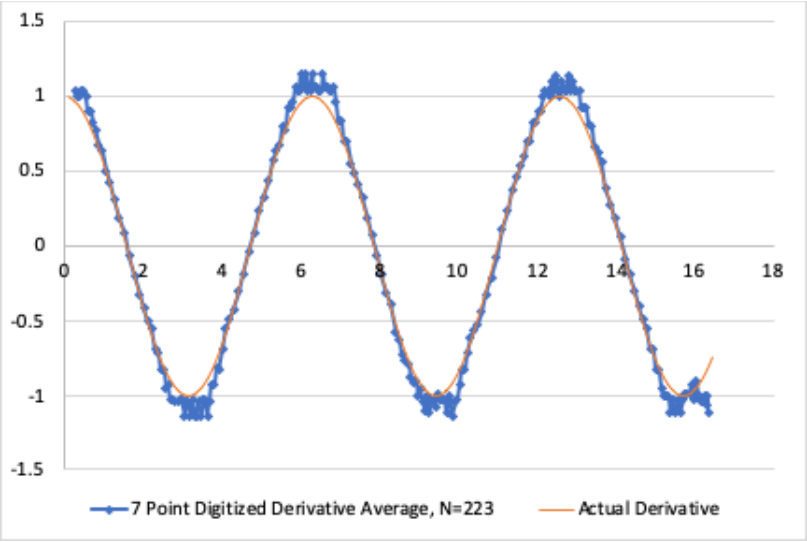


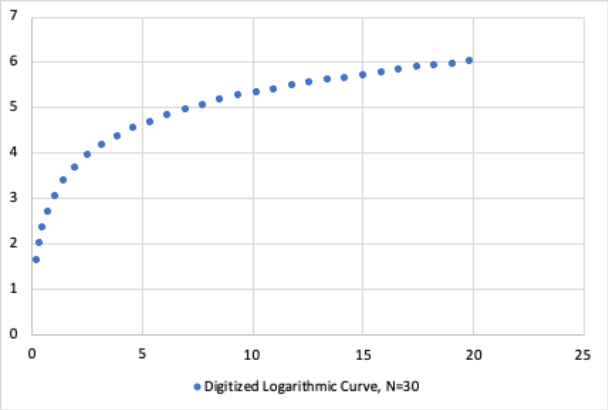
Figure 21. Seven-Point Digitized Derivative Average for N=223 on a sinusoidal curve

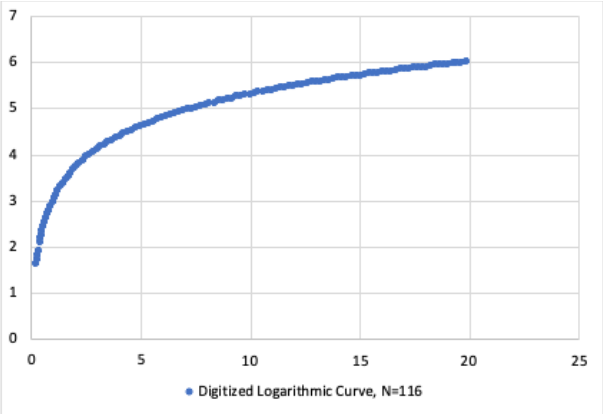
Based on these findings, curves that are digitized with a high N value can be averaged beyond the 5-point average to further minimize the fluctuations. While the number of points that averaged for any given curve, reaching a normalized sum square difference below 0.010 is accurate enough for this thesis. Therefore, for sinusoidal curves, three-point averages should be used for N>60, five-point averages for N values roughly between 60 and 200, and seven-point or greater averages for a N value greater than 200.

To further test this approach, the same process was applied to a logarithmic curve since its derivative gradually stagnates similar to most algae growth time curves. The equation $y=\ln(x)+3$ was chosen due to a simple derivative of $y'=1/x$, but was shifted up three so a larger part of the initial exponential portion of the graph is present within quadrant one and can be digitized. Point separation values of 20 and 5 pixels were chosen to yield N=30 and 116 points generated along the curve respectively. The data coordinates were imported from Engauge Digitizer and then the derivative and averages were calculated in the same manner as the sinusoidal curve, except for disabling the "Fill Corners" setting within the Segment Fill settings

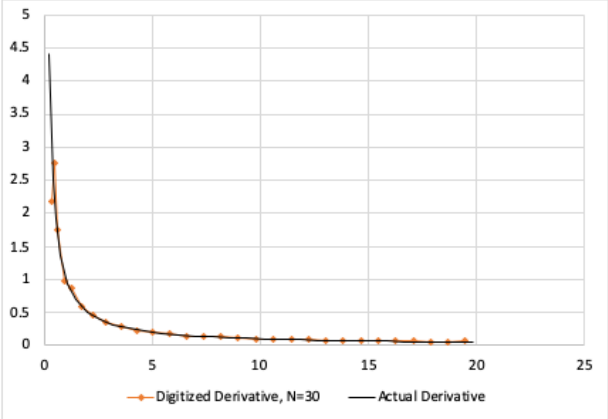
due to false detection of corners in the right side of the logarithmic curve. The logarithmic graph, its digitized derivatives, and 3 and 5-point averages to minimize fluctuations are shown in **Figure 22**.

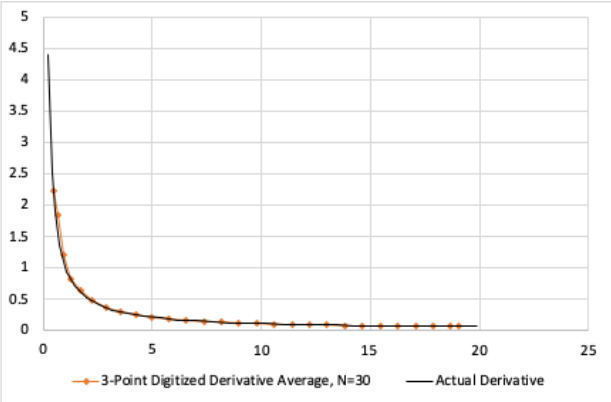
A.

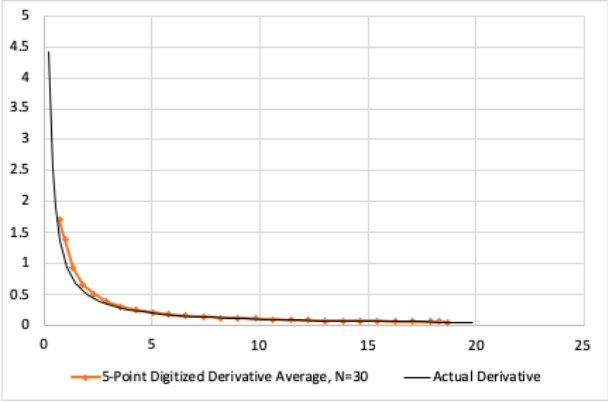




B.







C.

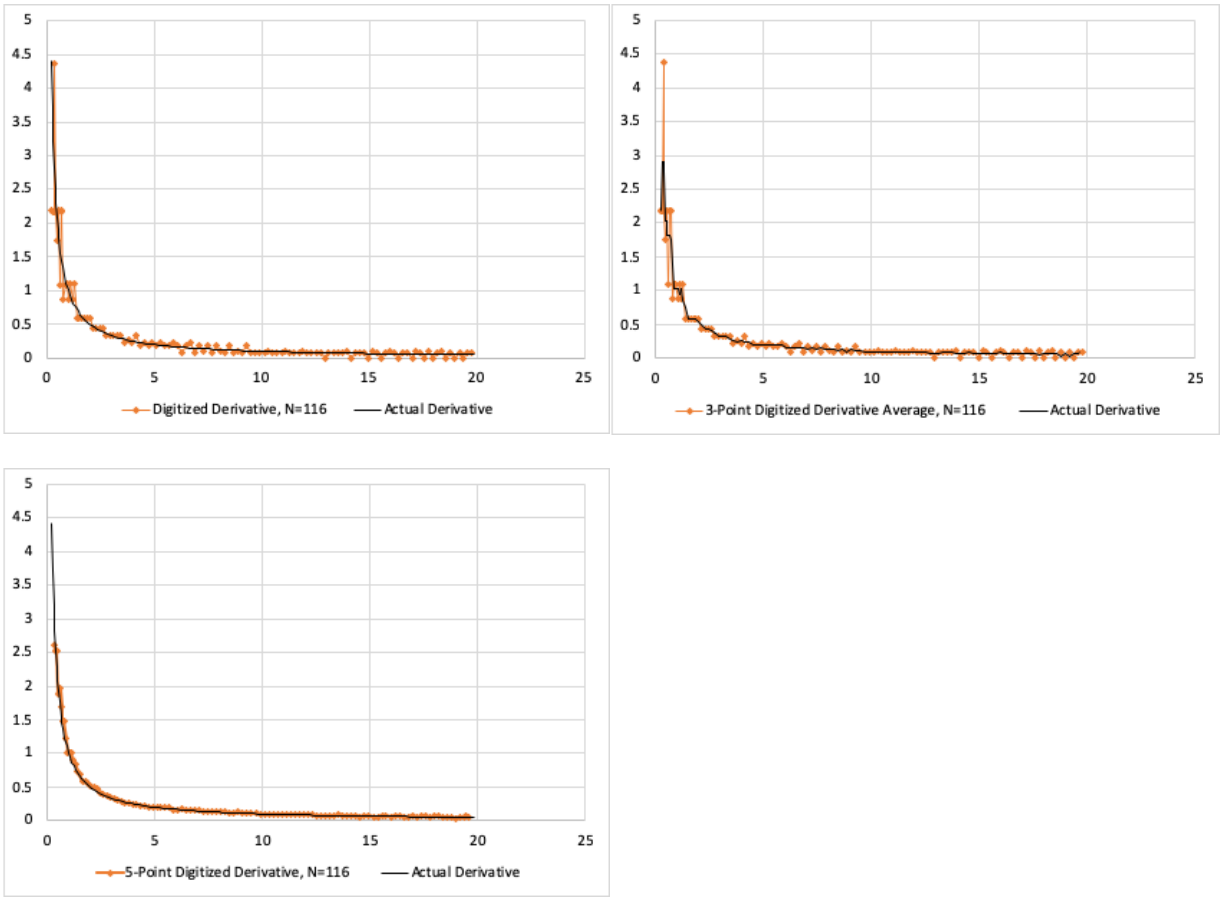


Figure 22. Digitized derivatives and 3 and 5-point derivative averages for the graph of y=ln(x)+3

As seen with the sinusoidal graph, the three-point digital derivative average fits the logarithmic curve better for low N values, while the five-point average shifts the digitized derivative slightly away from the actual derivative curve, as seen in section A of **Figure 23**. Furthermore, the five-point digitized derivative average visually minimizes the fluctuations in the N=116 curve in section B. To provide a quantitative index of which of the four alternatives best match the actual derivative, a sum square difference was taken between the digitized derivative values and the actual derivative values, shown in **Table 5**.

Table 5. Total and Normalized Sum square difference for the three and five-point average digitized derivatives for the graph y=ln(x)+3

Point Separation (Pixels)	20	5
Number of Points Generated Along the Graph (N)	30	116
Total Sum Square Difference (without averaging)	0.384	10.9
3-point average	0.130	2.18
5-point average	0.288	0.564
7-point average		0.350
Normalized Sum Square Difference (without averaging)	0.0132	0.0946
3-point average	0.00466	0.0192
5-point average	0.0106	0.00499
7-point average		0.00312

For the logarithmic graph, the digitized derivatives calculated with fewer points generated along the graph (the lower N value) have a lower normalized sum square difference for the three-point average than the five-point average while the higher N value derivative calculations have a lower normalized sum square difference with the five-point average. Unlike the sinusoidal graph, the lowest sum square difference arises from the normalized seven-point average with N=116, which means depending on the algae growth time curve that is digitized, either calculating derivatives from a digitized curve with a low N value with a three-point derivative average or a digitized curve with a high N value and subsequently 5 or 7-point average will yield digitized derivative values close to the actual derivative values. Since the fluctuations occur near the maximum derivative values, it is vital to minimize the fluctuations so the instantaneous productivity for experimental data is not over estimated.

Example Generation of Data Points along Experimental Data for Instantaneous Productivity Calculations

As mentioned previously, since the lines connecting the data points within Casadevall (1985)'s air-lift batch reactor hydrocarbon and biomass data are linear, a curve was hand-drawn based on visual best-fit and points were generated along the curve using Engauge Digitizer, as shown in **Figure 23**.

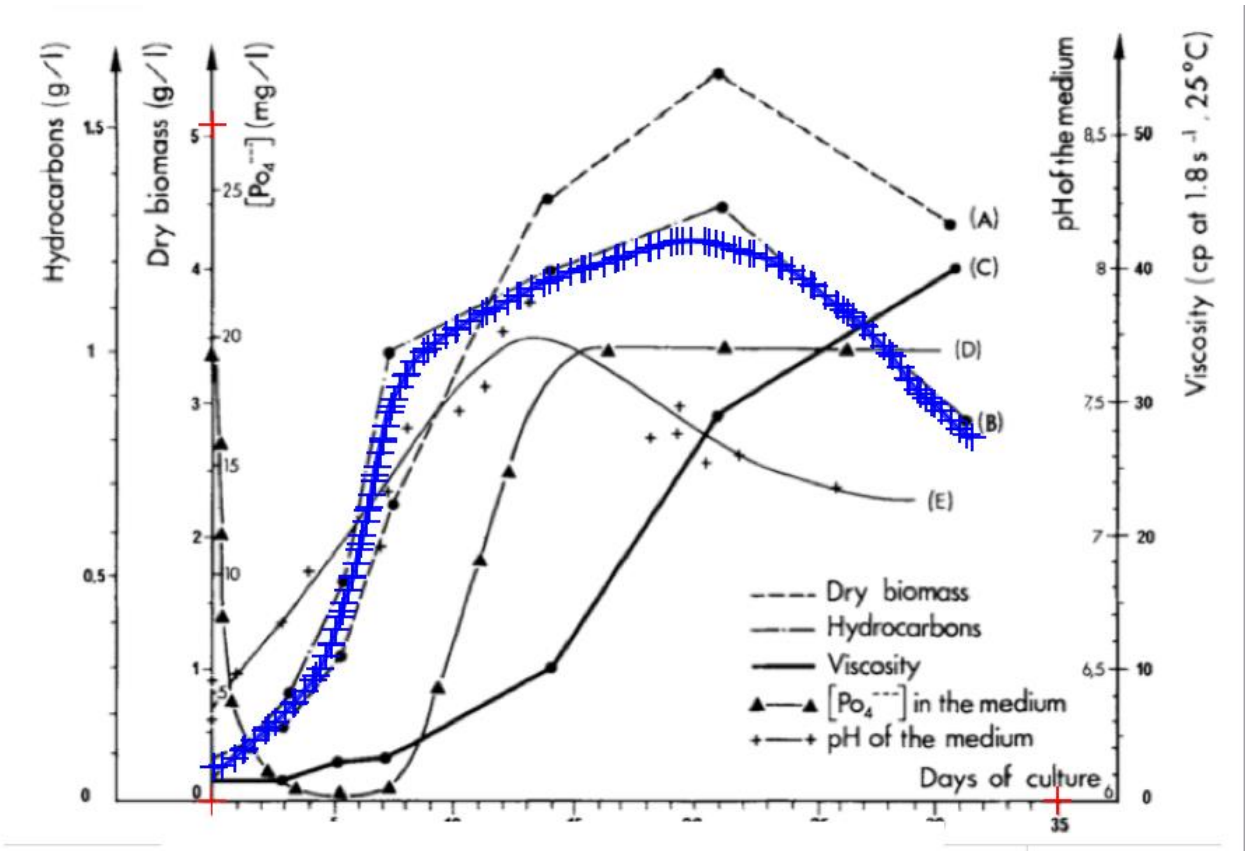


Figure 23. Points generated along the hand-drawn curve for hydrocarbon concentration in Casadevall (1985) air-lift batch reactor[17].

In **Figure 23**, 122 points were generated along the curve and were imported into Excel for further calculation described in Chapter 5.

Chapter 5

Example of Biomass and Hydrocarbon Productivity Calculations for Experimental *B. Braunii* Growth

Overview

This chapter combines the principles of productivity analysis described in Chapter 2 and 3 with the digitization of literature data discussed in Chapter 4 to undertake a specific example of starting with a literature report of hydrocarbon productivity to then calculate the productivity on a more meaningful basis for comparison to other reports of hydrocarbon productivity as part of a larger effort to establish a database of productivity for the unique hydrocarbon-production eukaryotic colony Botryococcus braunii being compiled at a wiki page, Botryococcus.org**.**

Methods

Engauge Digitizer, an application created by Mitchell et al[7], was used to convert graphical algae growth curves and other data into numerical data points in order to perform the productivity calculations. Notably this example included a table of the graphical reported values, so it is possible to compare the digitization to those values as a basis of validation of the method.

Experimental Set-Up

Casadevall et al conducted a growth experiment for *B. braunii* under both a batch and continuous process set-up to determine how physiological state, pH, phosphate levels affected *B. braunii*'s growth and hydrocarbon production.

Air-Lift Batch Growth of *B. braunii*

The growth conditions and reactor set-up for the *B. braunii* growth experiment by Casadevall et al (1985) is shown in **Table 6**. To note, since the cultures had continuous lighting during the experiment, a day was used as the measure of time rather than photohour so the productivity values could be compared to the reported values.

Table 6. Air-lift batch B. braunii experimental growth conditions in Casadevall et al (1985)[17]

Strain and Race	Temperature	Lighting Type	Light Flux Density (W/m^2)	Process Type	Light Cycle (light/dark)	Total Photohours of Growth	Reactor Volume
Strain 806/1: Race A	25°C	"Cold-white" Philips fluorescent bulbs alternated with "grow lux" Sylvania tubes	26	Batch	24/0 (continuous lighting)	504	1 L cylindrical tube, 65 mm diameter

Biomass and hydrocarbon concentration samples were taken over thirty-one days as shown by *B. braunii* growth curve **Figure 24**. Samples 1-5 were taken at the points indicated in **Figure 24**.

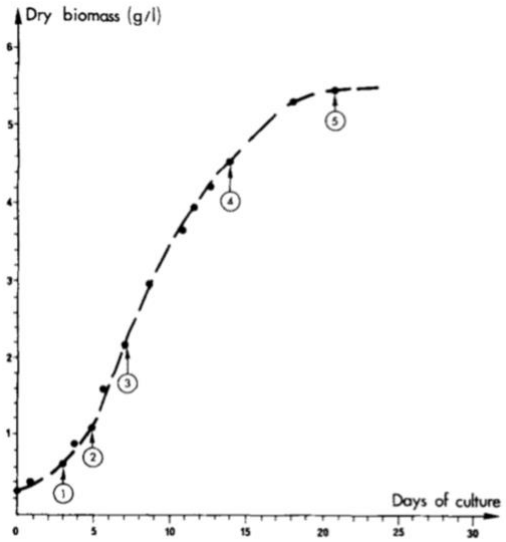


Figure 24. Biomass and hydrocarbon concentration sampling times for air-lift batch cultures (adapted from Casadevall et al (1985))

Using the digitization method to tabularize data described in Chapter 4, the hydrocarbon and biomass concentrations from the air-lift batch setup in **Figure 11** were digitized and the values are reported within **Table 7**. The literature reported hydrocarbon and biomass concentrations are given in **Table 8**.

Table 7. Digitized biomass and hydrocarbon concentrations for Casadevall (1985) air-lift batch cultures as well as percent error from literature reported concentrations

Day	Biomass Dry Weight (g/L)	Hydrocarbon Concentration (g/L)	% Error from Reported Biomass Concentration	% Error from Reported Hydrocarbon Concentration
0.0154	0.3286	0.04877	22	10.8
2.908	0.5556	0.23934	-4.2	8.8
5.3229	1.0902	0.48821	-3.52	14
7.4727	2.2342	0.99661	2.96	3.8
13.778	4.5341	1.17912	0.982	4.35
20.8933	5.4659	1.31925	0.292	3.88
31	4.3399	0.84579	2.84	5.7

Table 8. Reported biomass and hydrocarbon concentrations from air-lift batch cultures (adapted from Casadevall et al (1985))[17]

Sample[a]	External hydrocarbons (as % of total)	Nature[b] and relative abundance of major external hydrocarbons[c]			Total biomass dry weight (g/L)	Hydrocarbon yield of the culture (g/L)	Total hydrocarbon level (as % of biomass)	Hydrocarbon productivity[d]	
		$2\ \Delta C_{27}$	$2\ \Delta C_{29}$	$2\ \Delta C_{31}$				P_1 (g/L day)[e]	P_2 (g/day g of starting biomass)[f]
0	95.8	11.7	57.2	31.1	0.27	0.044	16.3		
1	98.8	12.5	56	31.5	0.58	0.22	37.9	0.059	0.218
2	96.3	10.8	65.5	23.7	1.13	0.43	38	0.105	0.181
3	99	11.2	62.8	26	2.17	0.96	44.2	0.265	0.234
4	96	13.3	62.7	24	4.49	1.13	25.2	0.025	0.011
5	93.7	10	58.9	31.1	5.45	1.27	23.3	0.020	0.0044
6	93.2	7	61.8	31.2	4.22	0.80	19	negative	negative
CB	95	17.5	56.25	26.25					

The percent errors between the digitized and actual biomass and hydrocarbon concentrations are included within **Table 7,** in which the sample numbers correspond to the sampling date given in **Figure 24**. Most of the errors are below five percent for the larger concentration values, but for lower values, the error reaches 22%. However, when comparing the

digitized and literature values, the error corresponds to a 0.0048 g/L difference, which is sufficiently low for insignificant impact on productivity calculations.

Since both digitized and actual biomass and hydrocarbon concentrations are available, productivity calculations were performed on both sets of data and compared to one another. Casadevall et al (1985) also reported their own calculated average volumetric and specific hydrocarbon productivities in **Table 8.** The reported average volumetric hydrocarbon productivity is calculated in the same way as described in Chapter 2. Casadevall (1985) calculated average specific hydrocarbon productivity using the initial biomass concentration over the period the productivity is calculated while this thesis calculates specific productivity with the final concentration for the calculation period to better account for the changing biomass concentration over time.

The average volumetric and specific productivities calculated from the reported hydrocarbon and biomass concentrations are included in **Table 9**, in which the days reported are digitized values from **Figure 24.** The average specific hydrocarbon productivity on the basis of photon is also reported and is calculated using the culture vessel volume and radius reported with Casadevall (1985) and summarized in **Table 6.** To obtain the lighted area of the vessel, height was found from the volume and diameter using the assumption of a right cylinder geometry, and then surface area was calculated assuming the entire vessel was lighted. Sample calculations for the average volumetric, specific (on the basis of day), and specific (on the basis of photon) are shown in **Equations 19 and 20** respectively.

$$AVHP = \frac{\Delta HC}{\Delta day} = \frac{(0.22-0.044)\,\frac{g\,HC}{L}}{(3.2-0)\,days} = 0.055\,\frac{\frac{gHC}{L}}{day} \qquad \textbf{Equation 19}$$

$$ASHP\ (day\ basis) = \frac{\frac{\Delta HC}{gDW_i}}{\Delta day} = \frac{\frac{(0.22-0.044)\,\frac{gHC}{L}}{0.58\,\frac{gDW}{L}}}{(3.2-0)\,days} = 0.0941\,\frac{\frac{gHC}{gDW}}{day} \qquad \textbf{Equation 20}$$

Table 9. Calculated Biomass and Hydrocarbon Productivity Values for Air-Lift Batch *B. braunii* batch experiment by Casadevall et al (1985) from Reported Hydrocarbon and Biomass Values

Days	Photohours	Average Volumetric Hydrocarbon Productivity (gHC/L/day)	Reported Average Hydrocarbon Volumetric Productivity (gHC/L/ day)	Average Specific Hydrocarbon Productivity (gHC/gDW/ day)	Reported Average Specific Hydrocarbon Productivity (gHC/gDW/ day)
0.0	0	N/A	N/A	N/A	N/A
3.2	77	0.055	0.059	0.0941	0.218
5.4	130	0.096	0.105	0.0849	0.181
7.3	175	0.282	0.265	0.130	0.234
14.0	336	0.025	0.025	0.0057	0.011
21.1	506	0.020	0.020	0.0036	0.0044
31.0	744	-0.047	negative	-0.0112	negative

The calculated and reported average volumetric hydrocarbon productivities are similar, the variations arising from the digitized day values since they were not explicitly reported within the literature. There is greater variation between the calculated and reported average specific hydrocarbon productivity, but that arises from the difference in specific productivity hydrocarbon calculations, in which the literature uses the initial biomass concentration of the two time points the samples are taken while the thesis uses the final biomass concentration. If the initial biomass concentration is used within the thesis calculations, the average specific hydrocarbon productivities are equivalent. Arguably using an average biomass level during this time period might be more appropriate, however, as digitization allows for measurements of instantaneous specific productivity, this time interval for cell density for such calculations becomes irrelevant.

The hydrocarbon productivity values calculated from the digitized data are given in **Table 10**, with the same productivity types calculated as reported in **Table 9.**

Table 10. Calculated Hydrocarbon Productivity Values for Air-Lift Batch *B. braunii* batch experiment by Casadevall et al (1985)

Day	Photohours	Average Volumetric Hydrocarbon Productivity (gHC/L/day)	Average Specific Hydrocarbon Productivity (gHC/gDW/ day) (initial biomass used)	Average Specific Hydrocarbon Productivity (gHC/gDW/ day) (final biomass used)	Average Specific Hydrocarbon Productivity (gHC/gDW/ photohour)
0.0154	0.370				
2.91	69.8	0.066	0.200	0.119	0.00494
5.32	127.7	0.103	0.185	0.0945	0.00394
7.47	179.3	0.236	0.217	0.106	0.00441
13.8	330.7	0.029	0.0130	0.00638	2.66E-04
20.9	501.4	0.0197	0.00434	0.00360	1.50E-04
31.0	744.0	-0.0468	-0.00857	-0.0108	-4.50E-04

The percent error between the productivity values calculated from the reported data, from digitized values, and the reported productivity values are given in **Table 11**.

Table 11. Percent differences between reported average specific hydrocarbon productivities and average specific hydrocarbon productivities (ASHP) calculated from reported and digitized data

Day	% Difference between Reported ASHP and ASHP from reported values (initial biomass)	% Difference between Reported ASHP and ASHP from reported values (final biomass)	% Difference between Reported ASHP and ASHP from digitized values (initial biomass used)	% Difference between Reported ASHP and ASHP from digitized values (final biomass used)	% Difference between ASHP from reported values and digitized values (initial biomass concentration used)	% Difference between ASHP from reported values and digitized values (final biomass concentration used)
0.0						
3.2	-7.25	-56.82	-8.03	-45.61	-0.84	25.98
5.4	-8.61	-53.09	2.48	-47.77	12.14	11.34
7.3	6.82	-44.38	-7.30	-54.77	-13.21	-18.68
14.0	6.33	-48.61	17.78	-41.96	10.76	12.93
21.1	-0.32	-17.88	-1.28	-18.11	-0.97	-0.29
31.0	N/A	N/A	N/A	N/A	-1.57	-4.01

Though the percent difference appears high for some of the comparisons, such as between the reported ASHP from the literature and the ASHP calculated using digitized data, the differences between the productivities are not very large when comparing productivities that both are calculated from the final biomass or initial biomass. The small differences in productivities are shown in **Figure 25,** which compares the average specific hydrocarbon productivities from the digitized and reported data compared to the reported productivity.

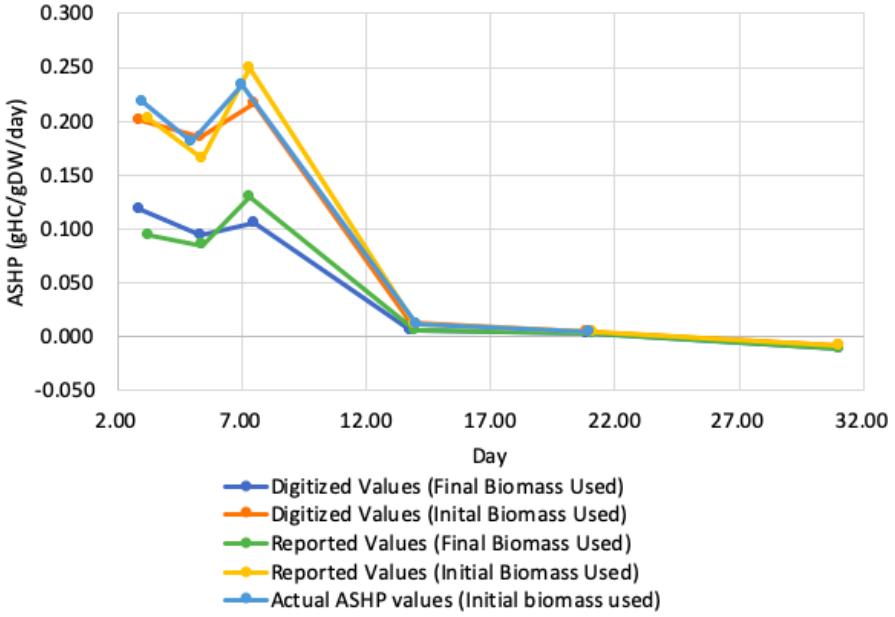


Figure 25. Average specific hydrocarbon productivity over time based on the different modes of data origin for *B. braunii* growth data from Casadevall (1985)

Given the similarity between the productivities calculated from the reported data and the digitized data, using digitized data for productivity calculations is a sufficiently accurate method when data is not tabulated in literature.

Calculation of Instantaneous Productivity for Casadevall (1985) Air-Lift Batch Culture Data

As described in Chapter 4, given the multiple hydrocarbon concentration points over time, a curve can be drawn between the points to determine instantaneous productivity. **Figure 26** shows the points generated along the curve using the Segment Fill tool in Engauge Digitizer based on **Figure 23**, with N=122.

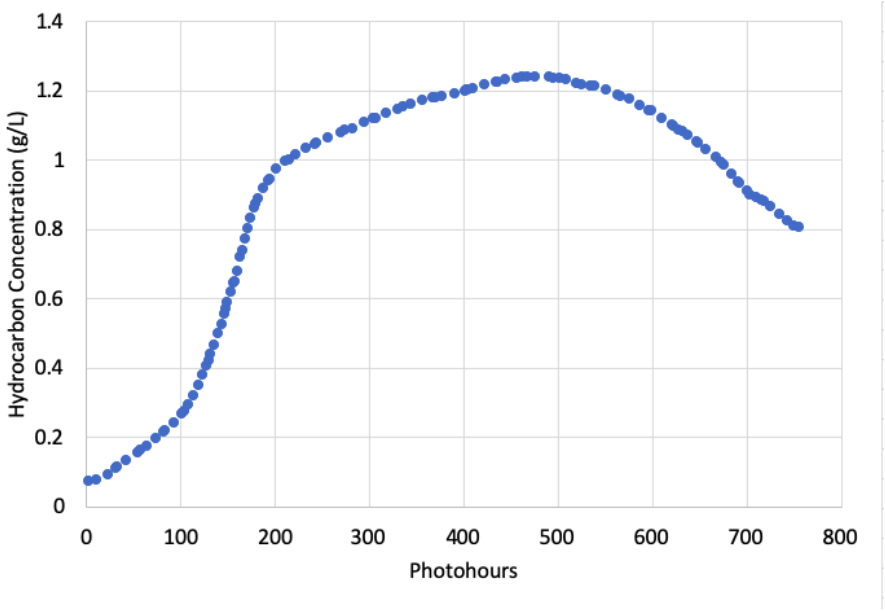


Figure 26. Digitized air-lift batch culture hydrocarbon accumulation graph from Casadevall (1985).

From these data points, the instantaneous volumetric hydrocarbon productivity was found using dy/dx between subsequent points, which yielded **Figure 27**.

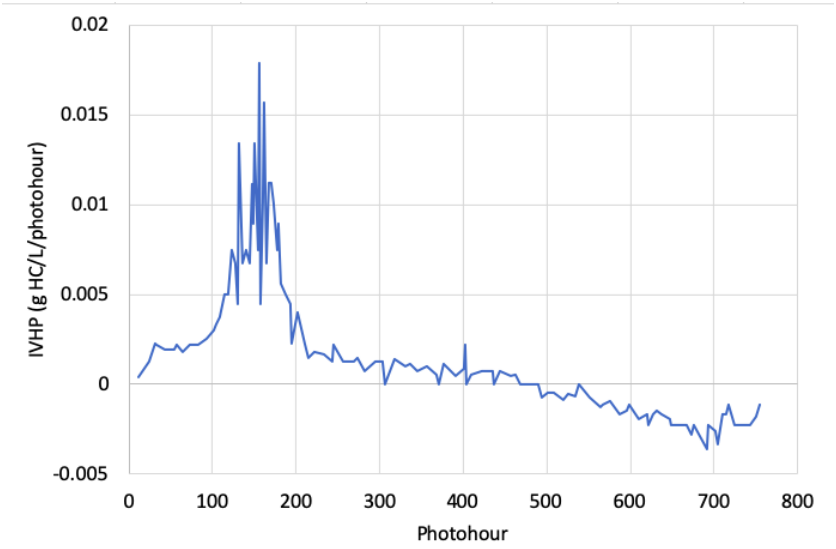


Figure 27. Instantaneous volumetric hydrocarbon productivity for Casadevall (1985) air-lift batch cultures.

As seen with the digitized derivative calculations in Chapter 4, **Figure 27** shows a lot of fluctuation in the derivative values, especially between 100 and 200 photohours. To minimize the fluctuations, a five-point derivative average was taken since N = 122, shown in **Figure 28**.

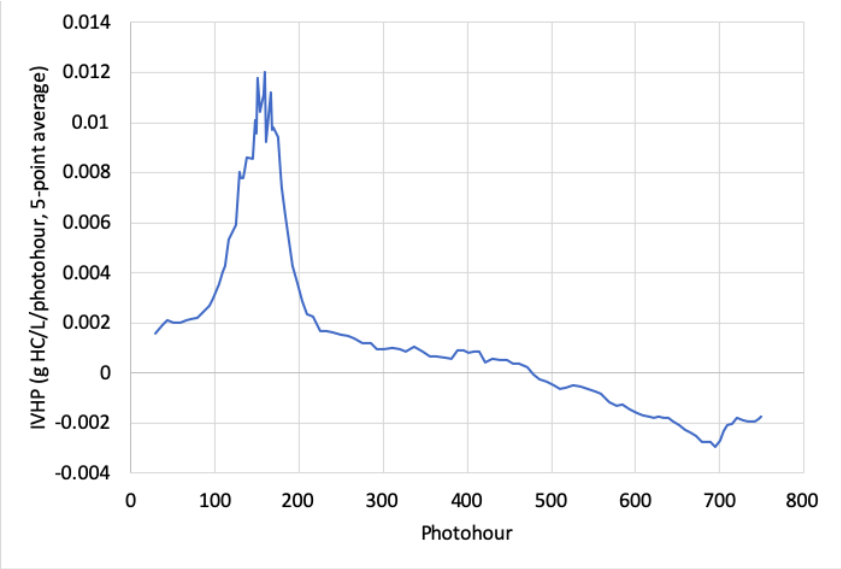


Figure 28. Five-point average taken for instantaneous volumetric hydrocarbon productivity of air-lift batch cultures from Casadevall (1985).

In comparison to **Figure 27,** the fluctuations are sufficiently suppressed within **Figure 28** and give a better indication of the maximum instantaneous volumetric hydrocarbon productivity, which is approximately 0.012 gHC/L/photohour.

Chapter 6

Key Conclusions and Future Work

Key Conclusions

Given the comparison with experimental tabulated data and digitized data from experimental growth curves, the data obtained from digitization yields similar productivity results as the productivities calculated literature reported hydrocarbon concentrations. As a result, for experimental data that is not explicitly tabulated, using digitization software can give insight into how the algae species and strain tested will roughly compare to another by normalizing the light given to the algae and then calculating the productivity. A conversion to photohours can provide a simple basis for different durations if diurnal or continuous light; however, the greatest challenge in making these comparisons is dealing with the interaction of lighting and the photobioreactor configuration. While the framework for such analyses-based growth yield for PAR photons is preliminarily outlined in this thesis, that analysis has not been conducted as we could not find a dataset with which to conduct that literature validation. The utility would be in the comparison of those calculations (with copious assumptions) for a breadth of studies. Future work should include digitizing the plethora of other hydrocarbon-yielding algae growth curves for comparison to illustrate the importance of promoting an analysis standard that can be compared to other studies. Only in this way can one start to assess which algae has the highest specific hydrocarbon productivity. Ideally, one could then then conduct further validation experiments that have more precise information about the growth and lighting

conditions as a basis for refining an assessment of which algae strains can lead towards the most productive capture of light energy into biofuels.

Productivity Calculations with Well-Defined Light Input

Productivity calculations on a photon-basis would be very insightful given how algae will almost certainly be grown under high-density, light-limited conditions within industrial settings. Therefore, reaching out to the researchers for culture vessel dimensions in order to calculate lighted area for the vessel would be the next step in these calculations in order to convert productivity on a photohour basis to photons, which indicates how efficient the algae is the light it is given as well as how much hydrocarbon it can produce with a single photon of light. In industrial growth of algae, light usage will be the limiting factor to hydrocarbon production, so if past experiments with algae can be compared on a light-usage basis, it unlocks a great deal of information regarding which algae should be used as a feedstock.

Establishing a method for performing productivity calculations can also help with future algae growth experiments, such as within bioreactors where the lighting is well-defined such as the "Alginator" in the Nugent thesis[8]. If algae growth experiments are run under similar conditions within the Alginator and then the hydrocarbon productivity of each is compared, then it allows for a more accurate comparison between algae species and strains due to similar growth conditions.

Open Source Productivity Calculations: Botryococcus.org

CurtisLab has created a website dedicated to open-access information about *B. braunii* growth and analysis in order to spur the conversion regarding algae biofuels. Current work is being done to digitize and calculate productivities for past algae growth experiments with *B. braunii*, which will ultimately be posted onto the website for public access. The website also allows for others to submit their growth data and analysis, which will ultimately allow for the compilation of a great deal of data in order to develop a more systematic approach to conducting assessments of which strain of *B. braunii* are most productive under different growth conditions with a particular focus on being able to translate that kinetic information to performance on a large scale.

BIBLIOGRAPHY

1. Tuerk, A. L. AN ASSESSMENT OF PHOTOSYNTHETIC BIOFUELS AND ELECTROFUELS TECHNOLOGIES UNDER RATE-LIMITED CONDITIONS. (Pennsylvania State University, 2011).
2. Masci, P.; Grognard, F.; Bernard, O. Microalgae biomass surface productivity optimization based on a photobioreactor model. *IFAC Proceedings Volumes* [Online] **2010**, *43* (6)**,** 180-185. https://sciencedirect.com (accessed March 10, 2021).
3. Eroglu, E.; Okada, S.; Melis, A. Hydrocarbon productivities in different Botryococcus strains: comparative methods in product quantification. *J. Appl. Phycol.* [Online] **2011,** *23*, 763-775. https://link.springer.com (Accessed March 8, 2021).
4. Myers, J. A.; Curtis, B. S. & Curtis, W. R. Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC Biophysics* [Online] **2013**, *6*(4). http://www.biomedcentral.com (accessed March 8, 2021).
5. Xu, L.; Liu, R.; Wang, F.; Liu, C. Development of a draft-tube airlift bioreactor for *Botryococcus braunii* with an optimized inner structure using computational fluid dynamics. *Bioresource Technology* [Online] **2012,** *119,* 300-305. https://sciencedirect.com (accessed March 10, 2021).
6. Kirk, J. T. O. Light capture by aquatic plants. In *Light and Photosynthesis in Aquatic Ecosystems,* 2nd ed.; Cambridge University Press, 1994; pp 252-270. https://doi.org/10.1017/CBO9780511623370.011
7. Mitchell, M.; Muftakhidinov, B.; Winchen, T. *et al. Engauge Digitizer Software,* version 12.1, 2019. http://markummitchell.github.io/engauge-digitizer.

8. Nugent, L. A. FIRST PRINCIPLES AIRLIFT PHOTOBIOREACTOR FOR THE ENERGY BALANCE OF MICROALGAE FOR BIOFUELS APPLICATIONS. (Pennsylvania State University, 2020).
9. Ritchie, H.; Roser, M. CO_2 and Greenhouse Gas Emissions, 2020. Our World in Data. https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions (accessed March 31, 2021). Data from Data supplement to the Global Carbon Project 2020. Integrated Carbon Observation System. https://doi.org/10.18160/gcp-2020
10. Khan, M. I.; Shin, J. H.; Kim, J. D. The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. *Microb.Cell Fact.* [Online], **2018,** *17*(1), 36. https://www.ncbi.nlm.nih.gov (accessed March 31, 2021).
11. Ratha, S.K.; Rao, P.H.; Govindaswamy, K. *et al.* A rapid and reliable method for estimating microalgal biomass using a moisture analyser. *J Appl Phycol* [Online] **2016**, *28***,** 1725–1734. https://link.springer.com (accessed March 31, 2021).
12. Grady, L. K. A BIOPROCESSING COMPARISON OF HIGH DENSITY BOTRYOCOCCUS BRAUNII AND CHLORELLA VULGARIS VERIFYING LIGHT LIMITED GROWTH. (The Pennsylvania State University, 2010).
13. Rao, R. A.; Sarada, R.; Ravishankar, G. A.; Phang, S. M. Industrial Production of Microalgal Cell-Mass and Bioactive Constituents from Green Microalga-Botryococcus braunii. In *Recent Advances in Microalgal Biotechnology*, Liu, J., Sun, Z., Gerken, H., Eds; OMICS Group eBooks, 2014; pp 1-19. https://www.researchgate.net (accessed March 31, 2021).

14. NSF-Small Business Innovation Research (SBIR) Algae Screen Test, Award #0945592. December 26, 2010.

15. Boonma, S.; Vacharapiyasophon, P.; Peerapornpisal, Y.; Pekkoh, J.; Pumas, C. Isolation and Cultivation of Botryococcus braunii Kützing from Northern Thailand. *Chiang Mai J. Sci.*[Online], **2014**; *41*(2), 298-306. http://epg.science.cmu.ac.th/ejournal (accessed April 5, 2021).

16. Farag, I.; Price, K. Resources Conservation in Microalgae Biodiesel Production. *International Journal of Engineering and Technical Research* [Online], **2013,** *1*(8), 45-56. http://www.erpublication.org (accessed April 6, 2021).

17. Casadevall, E.; Dif, D.; Largeau, C, *et al.* Studies on Batch and Continuous Cultures of *Botryococcus braunii*: Hydrocaron Production in Relation to Physiological State, Cell Ultrastructure, and Phosphate Nutrition. *Biotechnol. Bioeng.* [Online] **1985,** *27*(3), 286-295. https://onlinelibrary.wiley.com (accessed April 16, 2021).

ACADEMIC VITA

Hinkal Patel

EDUCATION

The Pennsylvania State University, Schreyer Honors College. Graduation: August 2021
Bachelor of Science in Chemical Engineering, Minor in Chemistry

WORK EXPERIENCE

Purification Development Co-op *June 2020—December 2020*
Regeneron Pharmaceuticals, Tarrytown, NY

- Developed a web application to streamline the employee verification process to compare virtual inventory within a database and the physical lab inventory to save time when verifying as well as preventing consumable waste.
- Incorporated employee selection features into the app user interface to allow users to access lab consumable information from multiple databases into the app
- Maintained data integrity while coding through version control with GIT and Bitbucket.

RESEARCH EXPERIENCE

Member of Curtis Research Lab, Penn State *Aug 2017—Present*
Curtis Lab Research Assistant *May 2018—July 2018*

- Conducted productivity analysis by photosynthetic photon flux density for *B. braunii* algae hydrocarbons in relation to biofuels.
- Adept at lab techniques such as sterile technique, subculturing plants and algae, and PCR. Maintained tissue culture lines of tomato, cacao, and yam as well as axenic colonies of algae. Experienced with dilution plating and algae streaking to isolate a single algae colony. Also helped maintain a tomato genetic transformation pipeline.
- Ran study of differential protein expression of tomatoes under drought stress and analyzed significant water loss under drought stress using Analysis of Variance
- Coded a calculator to scale and compare media recipes to facilitate media elemental composition analysis.

LEADERSHIP

K-12 Outreach Visit Chair, Society of Women Engineers *Sept 2017—May 2019*

- Organized and led the first SWE STEM Expo at Penn State for about fifty elementary students to spark interest in STEM .
- Established contact and collaborated with thirty Penn State clubs and organizations to present at the SWE STEM Expo.

HONORS

2020 CSL Behring Innovation Scholarship