Evaluation of a Benchmark Model of Microalgae productivity towards Global Implementation

Guerbine Fils-aime

Dr. Zackary Johnson, Advisor

April 10, 2020

Master of Environmental Management from Nicholas School of the Environment at Duke University

I certify the following:

1. Does this proposed MP involve human subjects research? ___ Yes __x__No
   a. If yes, has an approved IRB protocol been obtained? ___ Yes ____No

2. Does this proposed MP involve the use of animals in research? ___ Yes __x__No

   a. If yes, has an approved IACUC protocol been obtained? ___ Yes ____No

3. Does this proposed MP involve signing a non-disclosure agreement? ___ Yes __x__No

   a. If yes, does the advisor have a signed copy? ___ Yes ____No

Student Signature:_______________________________ Date:_____________

Advisor Signature:________________________________ Date:_____________

Introduction

The growth of the human population and increasing demands on resources has led to a large increase of CO_2 emissions, which have been causing negative changes in climate. Models of future climate trends predict that the consequences will worsen if the emissions continue (or increase) at the current pace (Karl & Trenberth, 2003; Oreskes, 2004; UN, 2017; NOAA, 2017a; NOAA, 2017b, Beal et al., 2018). Even if humans stop the shift of carbon dioxide into the atmosphere relatively quickly, there will be a prolonged period before the effects of climbing CO_2 levels stabilize. If there is any chance for a stable climate and environment in the future, there is a need for zero carbon and net negative carbon emission technologies which could aid the effort to decrease atmospheric carbon dioxide levels. These technologies include solar, wind, and nuclear power, fossil energy with carbon capture and storage [CCS], carbon neutral crops, afforestation, and bioenergy with CCS (BECCS) or algae BECCS (Fuss et al., 2014; Greene et al., 2010, 2017; IPCC, 2013; Walsh et al., 2017; Beal et al., 2018).

Of these technologies, algal production bioenergy and carbon capture storage (ABECCS) is particularly promising because it can generate high amounts of energy, reduce CO_2 emissions naturally through photosynthetic processes, and simultaneously help to fulfil food production needs (Greene et al., 2016; Beal et al., 2018). This final component is particularly important because there is an increase in demand for food production and greenhouse emission reduction, however there is a concern that industries cannot keep up with food production while also maintaining greenhouse reduction goals (Walsh et al. 2016). ABECCS is a unique system compared to other BECCS because while BECCS offer help to mitigate climate issues and provide new resources for energy, they often combat food production space and needs (Greene et al., 2016). ABECCS does not conflict with space for food production most times-- algae fuel production contributes to food demands with high yield and has the ability to grow on non arable land that food crops cannot grow on anyways (Moody et al., 2014, Walsh et al., 2016; Greene et al., 2016). All aspects of algae biofuel production can help us reach food and energy UN established environmental goals, making it promising technology. (Walsh et al. 2016).

At the current stage of technology, the ability to perform all of these goals on a large scale is prohibitively costly (Beal et al 2015, 2018). The high costs to implement algae biofuel production prevents it from becoming widespread (Loftus & Johnson 2017). There is also a need to expand on how to increase efficiency for ABECCS to be competitive on the capitalist market (Beal et al 2015, 2018). The high costs are the result of space (land) needs, costs of water necessary, capital and labor expenses, energy use, and chemical and fertilizing additives for the algae, among various other factors in the field that may drive others away from this promising market.

While these factors represent significant barriers to implementation, they could be overcome if algae productivity were greater. However, the poor understanding of what drives productivity of ABECCS, and how to improve it, is a significant barrier to make it more cost competitive at this time. Many factors can influence the productivity and ultimate yield to maximize ABECCS efficiency and profitability such as water reuse, optimal temperature, and strain selection. Calculating ways to predict an increase in ABECCS yield is important to its broader implementation as an algal and forest feed and fuel eco-system. This system holds great promise but productivity is the proximal barrier to implementation. Because large scale experimentation is resource and time intensive, accurate models of productivity are essential to predict and ultimately enhance algal growth and yield towards making ABECCs become a competitive option. Towards this goal, Huntley et al (2015) have developed a model for algae (biofuel) productivity based on some key factors such as total nitrogen (TN), and particulate organic carbon (POC), to predict expected algae biomass productivity. This study seeks to test the model's fidelity with measured values thus assessing its usefulness in the broader algae field. This assessment can also lead to improvements to the model.

.Methods

Work began first with summer research at the Duke University Marine Lab in Beaufort, North Carolina. Gathering quantitative and qualitative data, harvesting algae biomass, understanding how information is inputted into the Marine Algae Industrial Consortium (MAGIC) database and where to find this information was conducted during that time. The main goals of the MAGIC project is to assess the productivity potential of algae strains, to find new ways to culture the algae efficiently, to develop two harvesting methods to gain both biofuels and food ingredients, demonstrate productivity of algae cultivation and to show the commercialization potential of this market.

The main methods for the project is data processing from a large bank of data recorded from MAGIC, which has a primary algae growth field site in Beaufort, NC. These data range from environmental, processed (dry weight, oxygen levels, etc), and output production data. Modelling methods were adapted from Huntley et al. (2015). These calculated predicted data should approximate the corresponding existing productivity data if the model acts as it should as an estimation tool.

The model used held three major assumptions which allowed the predicted outputs to be accurate estimates. These assumptions concern the DIN/PON, C:N ratio, and the nutrients utilized in the following steps. Dissolved inorganic nitrogen (DIN) values were normalized by particulate organic nitrogen (PON); total nitrogen (TN) was calculated by adding DIN and PON values. For the best growth in the algae, DIN/PON is ~2 (Figure 1).

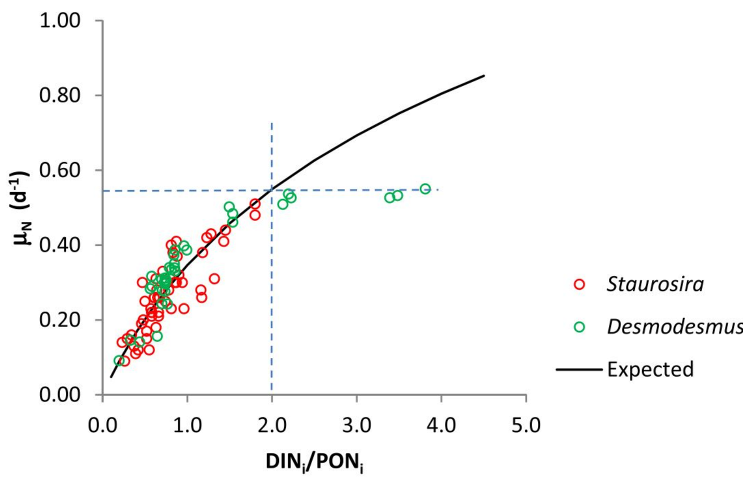

Figure 1. Assumption one, DIN/PON shows the best growth done during cultivation. This graph was from Huntley et al's demonstration of optimal growth (after Huntley et al. 2015).

The next main assumption is that the ratio of carbon to nitrogen (C:N) for the algae is about 7.7 mol:mol. The C:N ratio was calculated for both the final time points and initial time points of production by dividing particulates organic carbon (POC, µM) by PON. Then the gross Carbon value in the algae was calculated by multiplying the TN by the final C:N ratio. The initial value was found by multiplying initial PON by initial C:N.

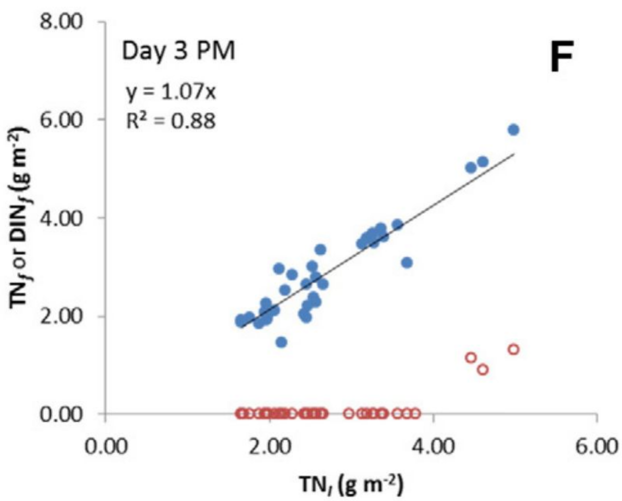

Figure 2. The algae from Huntley et al (2015) used all nutrients provided, which was an assumption established for the model. The total Nitrogen is almost completely depleted by the third day (after Huntley et al. 2015).

<u>The third major assumption is the algae utilizes all the nutrients provided when productivity is calculated</u> (Figure 2). The net growth was calculated by subtracting gross and initial C amounts, this compares initial C and final C gross values added to the pond. Using the third assumption, the nutrients used allows the model to estimate net growth. Once these values are found, the model can then calculate g DW/m2/day productivity from the net growth divided by the POC rate of change, then multiplied by dry weight (DW) rate of change, and then finally dividing by the number of days between measurements. For this paper, the rates of change provided by Huntley et al (2015) were used during the modeling (POC = 0.498, DW = 1.028). Then to determine if the model is accurate, the calculated biomass is compared to g DW/m2/day values recorded in the MAGIC database.

This was performed for three strains of algae recorded for the MAGIC data, H1117, C046, and S002. The model should approximately predict biomass data accounting for correct assumptions and values regarding nitrogen, carbon, etc. This model could be used to accurately determine the best way to maximize the algae productivity and efficiency if proven to have statistically similar values predicted.

All of the assumptions from the Huntley et al (2015) paper were assumed for this evaluation. They were recorded through DIN values normalized by PON values for every data entry, carbon values compared to nitrogen values for every entry to find the ratio, and the nutrient assumption was taken as pre-determined.

Microsoft Excel and R studio were used to process this data. After the modelling is done for the MAGIC sites, statistics were performed to determine the fidelity of the model predictions with observed numbers. This evaluation will give us the idea of how well this model (Huntley et al (2015)) predicts the algae productivity.

Results

During calculations, any recorded data for nutrients that were not available (N/A) and resulted in an 'undefined' or 'error' values were omitted from final statistics and graphing.

The model calculations gave higher predicted biomass/productivity than the actual observed values recorded in MAGIC database. If the model is used as it currently is formatted, it will overestimate values of biomass than what will actually be achieved (Table 1.)

Table 1. The observed values for each algae strain from the MAGIC database and the corresponding calculated values predicted by the model. All N/A values were removed from the final table.

H1117				C046				S002	
Observed Values	Predicted Values	Observed Values	Predicted Values	Observed Values	Predicted values	Observed Values	Predicted Values	Observed Values	Predicted Values
4.03319	14.3104038273679	3.968534	9.07625191713372	5.244828	29.7290005474494	6.713793	33.4114640559068	7.052586	30.1982414830839
11.630172	9.34084072156671	5.816379	17.3721466427514	1.52069	31.6466490473391	8.118103	32.7284566195589	3.332328	15.8931064885469
27.463966	19.2421958586017	7.584052	47.1445816035066	2.133621	19.618938111384	8.650862	15.7440397295207	5.689655	48.5729620627265
8.757931	15.7677743565957	8.428448	28.7791061888663	6.010345	35.5304872496736	2.614655	32.9185253920145	4.473276	21.5974907496628
-7.176724	9.0122585563981	12.409914	8.51878372954674	4.357759	12.3896008246242	2.438793	14.3473402698147	9.369828	49.9575724700515
4.50931	28.7255352202601	7.98556	29.1143821528204	3.32069	40.5774768633417	4.567241	30.7126494900715	9.289655	9.64122136905779
8.596552	9.07785909554435	6.590948	16.3657084676765	3.537931	36.227712010714	3.131897	30.398741431533	9.733675	7.29547079559334
13.712069	10.4097863547911	8.156897	24.3496390472354	3.15	17.5148700276124	6.943966	16.6923112368345		
12.805049	27.9678707519242	6.956897	9.87714223959353	8.896552	33.1695442653777	4.939655	32.2198865484372		
9.653276	19.2477253412906	4.635776	28.5801371600707	3.700862	13.7302934622494	3.331034	8.43433662132163		
7.570474	9.22044430070836	5.79181	14.1045237621835	6.956897	30.2657474269461	3.15194	28.5886086339836		
3.137862	27.3834152209768	5.327586	31.591568334219	5.581034	32.4925136974558	4.38944	10.1799539773675		
12.219517	45.2646813311867	4.952586	10.1732401962216	10.276293	18.1619523880894	5.482759	10.1369069227348		
2.335345	15.1663721540289	6.387284	30.4439335604772	7.016379	39.3980711194035	5.827833	18.4874769064564		
-0.036853	12.7523966289484	3.440948	14.8537047799301	8.746552	16.144981169837	6.850862	18.4543061749994		
11.092241	26.8011935604372	6.408621	28.8126794099328	12.181034	31.1467612700103	6.730603	12.4263359378835		
5.389655	20.1887903748326	11.987069	14.3493149161463	5.11681	15.1668975695457	5.787931	16.1646932584297		
9.92069	29.0521200648771			5.940517	38.718409967722	4.94069	15.9091549449637		
7.875	9.27464736967426			6.351437	17.3582646595608	-2.791164	10.5880427157996		
12.601149	46.4276586325448			8.601724	32.2241604856927	3.080172	16.6054080073631		
7.272931	18.1161538550713			4.735345	32.6076455840382	5.151724	9.34072957865599		
3.662931	30.9451791453022			5.725862	16.9210083549525	1.628017	31.3428547085764		
10.525862	10.0900551117573			6.460345	34.2788627638549	-0.640086	29.7785111376982		
8.480172	30.3539577018557			5.724425	16.6170300745289	2.424569	8.43433020570999		
11.405172	19.1869202634372			5.731034	30.7017573307132				
2.886207	29.70651041047			4.287931	33.5546914761746				
6.874138	12.494833878876			1.753448	16.8055922235349				
3.65431	15.1305493952841			12.356897	34.8873688146783				
6.901293	40.303848478403			10.187069	10.7493585123069				
14.834483	34.3741958074754			7.799353	37.9801396690066				
9.205603	9.77224070422662			6.309698	23.0391073334933				
7.411422	35.7394638182807			8.221552	31.0820997866984				
2.315948	9.65863968617948			3.561207	15.1221516664263				

A paired T-Test was ran to find if these values were significantly different. The alpha value was 0.05 for all two-tailed t-test and equal variance was considered false. The observed data for the strain data was the first array and the calculated data was the second array in the t-test. For the first algae strain (H1117), the observed values recorded and the theoretical calculated values from the model were statistically different (alpha: 0.05, 0.001; p-value:

2.06E-11). For the second strain (C046), the two values were also significantly different (alpha: 0.05, 0.001; p-value: 1.97E-20). The third strain (S002) did not have enough values to test for significance.

The observed and calculated values were graphed against each other to see how well the regression connection was between the two factors. For H1117, the fit was low (R^2 = 0.025, Fig 3). For C046 strain, the fit was also low (R^2 = 0.03, Fig 4). For S002, the fit was low but did not have enough values for it to be of significant importance.

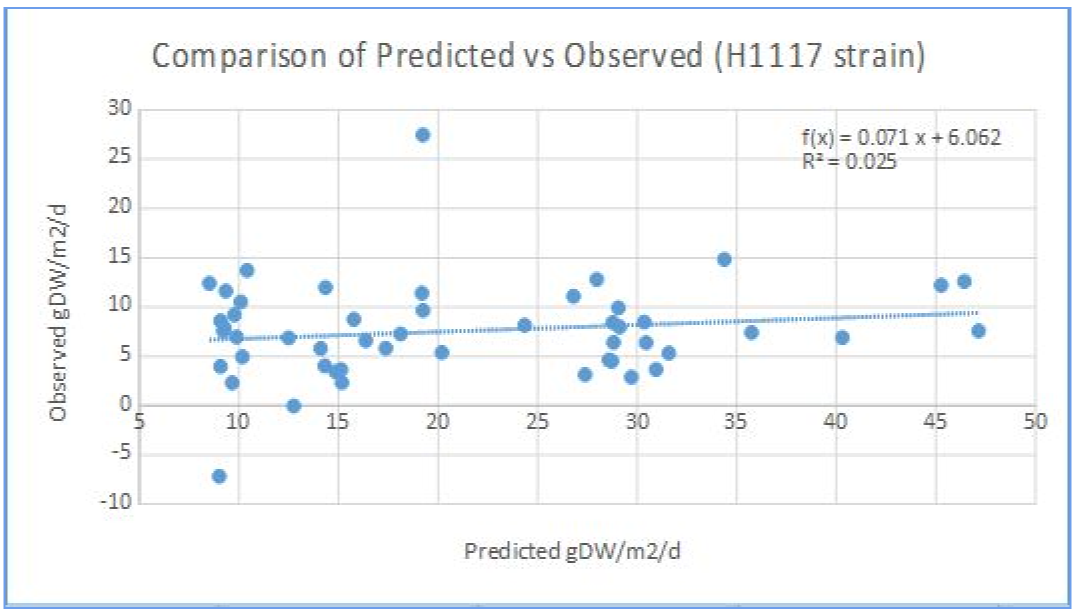

Figure 3. MAGIC Algae strain H1117 comparison of the observed DW/m2/day from existing data versus the predicted productivity given by the model, and the correlation (R^2 = 0.025) is poor.

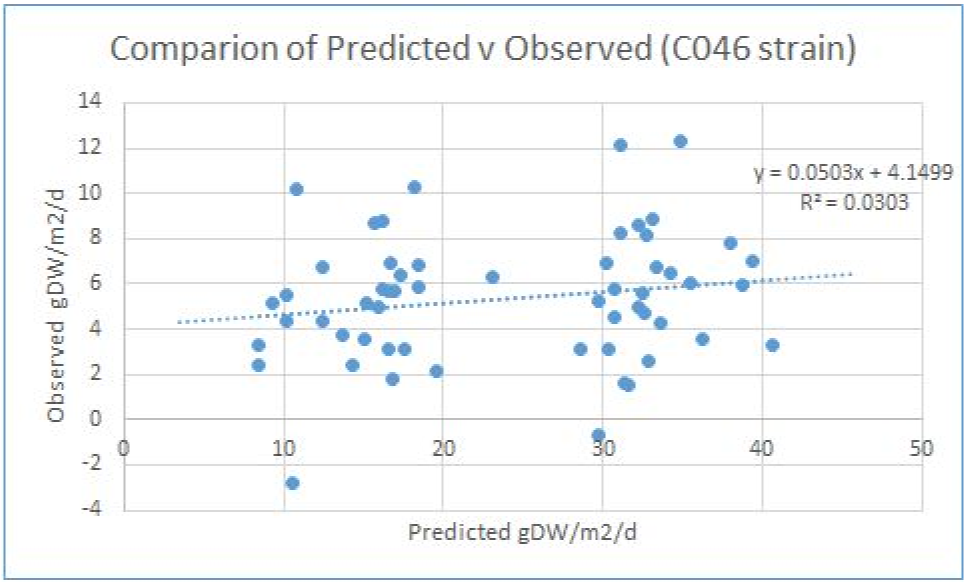

Figure 4. MAGIC Algae strain C046 comparison of the observed DW/m2/day from existing data versus the predicted productivity given by the model, and the correlation (R^2 = 0.030) is poor.

In an effort to try and isolate any driving factors that makes the model overestimate, H1117 was used as a representative example of how abiotic factors influenced the biomass (assuming it would be similar to the other strains). We can see in the below graphs that the model is over estimating values despite the abiotic factors. The model did not demonstrate an optimal growth level based on pH, but overestimated values across the recorded pH values (Figure 5). There is no apparent pattern of increasing temperature resulting in estimated increased algae production (Figure 6). And there is no obvious association with increasing recieved sunlight with estimated increase in net growth (Figure 7).

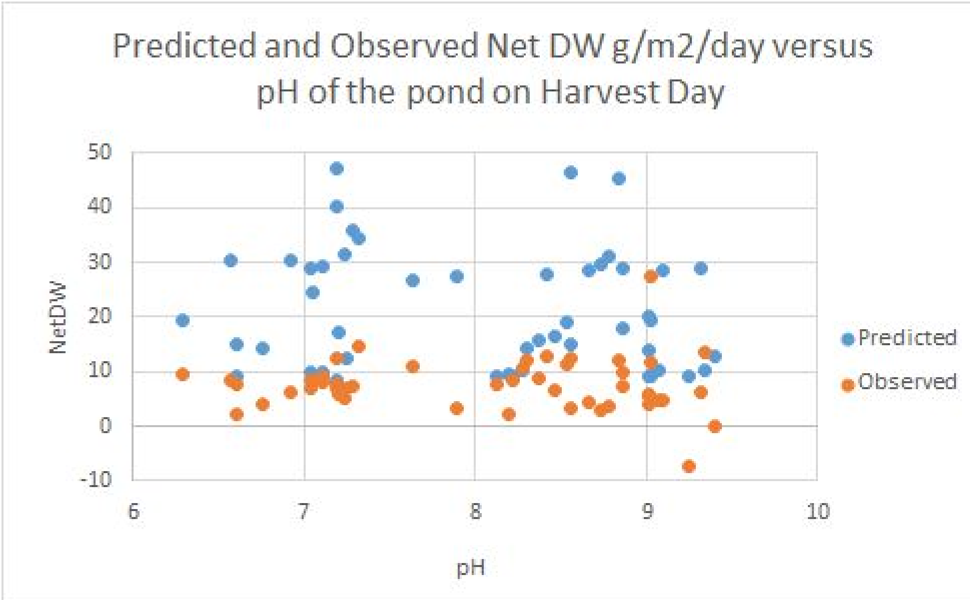

Figure 5. Model values and Observed values graphed against pH of the water throughout records for strain H1117.

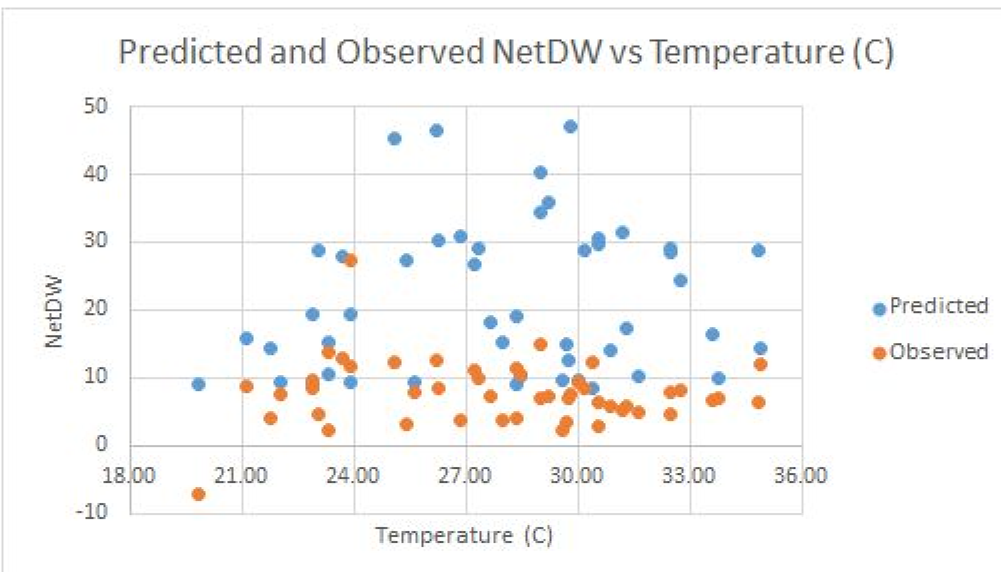

Figure 6. Model values and Observed values graphed against temperature (C) of the water throughout records for strain H1117.

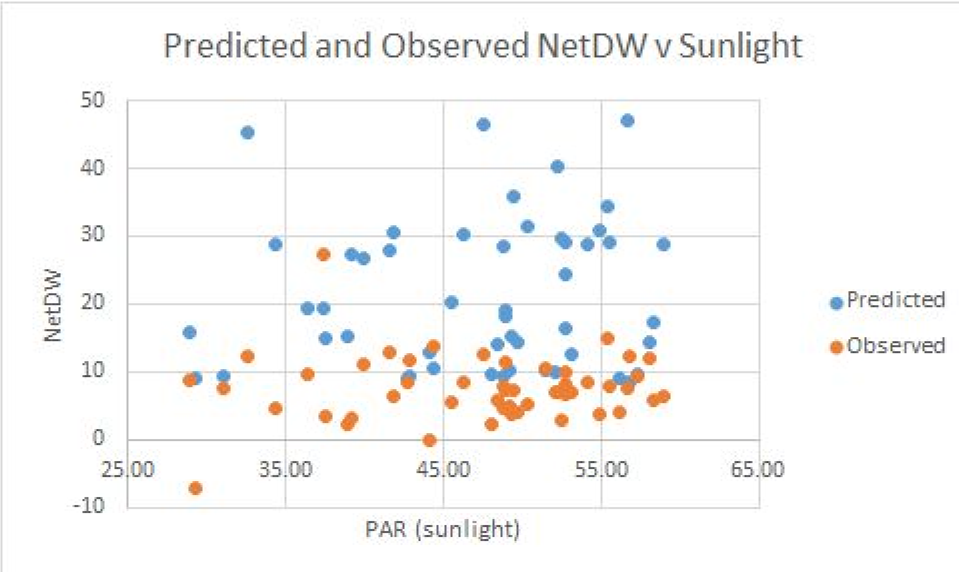

Figure 7. Model values and Observed values graphed against different levels of sunlight received throughout records for strain H1117.

Discussion

The results indicated that the model outlined in Huntley et al. (2015) does not give good approximations for predicting algae growth and productivity. The model does not predict proper algae biomass/productivity based on MAGIC dataset. It overestimates the values as far as can be seen on all aspects even when accounting for abiotic factors such as temperature and sunlight.

The fact that the two values were statistically different, even more than alpha of 0.01, means the model does not give a good approximation of the algae productivity. If the numbers appeared different but were not statistically different, it would still be a good model to use as an estimation of what you can get out of your resources.

A linear graph where R^2 is higher than 0.8, or mostly a straight line slope, we would see that the model can estimate the observed values well in a one-to-one agreement in values. But since it is poorly fitting, we see the estimations are not very well correlated. All these tests show that the Huntley et al. model should be improved or updated to be a proper tool in the new emerging field of algae production.

The data does not fit well because the data violated the necessary assumptions. The first assumption was that maximum algae growth would be DIN/PON equals 2, but instead, the values fluctuated on a broad range. This indicates some strains grew better on certain days than others in reality and the model did not account for this. The next assumption was that the algae uses all the nutrients provided, but the real data indicates that there were different rates of nutrients usage. This was another violation in the model assumptions that led to over estimations. The next assumption was that net growth could be used to estimate biomass, which was not violated in these processes. This paper also used the conversion slopes from Huntley et al., because calculated rates of change resulted in negative and incorrect values. Compared to Huntlet et al. (2015) where it worked, most assumptions were met with a DIN/PON approximating 2, nutrient usage was maxed at the values used, and their slopes were those they calculated based on the data they had on hand. All of these reasons may have been why the model overestimated values in most cases.

Other issues with the model might have been human error in calculations or data entry. This was also a test of an algae prediction model on one type of database, MAGIC. It could be of interest to see if the model consistently fails with other algae growing databases to accurately assess the model's validity. MAGIC is a relatively new project with limited data, such as what was seen with S002 data that had few recorded points at the time of calculations. There is also a need for consistent measurements of nutrients and other factors so that there are less N/A values in the calculations and more values could be considered.

Conclusion

We see that the algae model as described by Huntley et al. (2015) is not able to be applied more broadly to other datasets that need predictive biomass values. In Huntley et al. (2015), once they were able to satisfy the needed assumptions, they matched the calculated biomass with observed values. Other algae project datasets do not always have the necessary assumptions, or the data recorded in the necessary way to get approximate calculated biomass productivity values. This defeats the purpose of the model to be applied in a broader sense if violating the assumptions lead to overestimated values. The model will need to be improved to prevent a large skew of estimation from the assumptions being violated and to a system that can account for slight variation between data records.

This project is important to determine the validity of an algae model. The algae fuel industry is an emerging area of commerce that holds promise in the future of transportation and food production. It is a net negative carbon emission technology that aids in the fight of elevated carbon in the atmosphere that contributes to climate change, so it is important to have new and improving tools that allows the industry to become competitive with fossil and current fuels. This model was made to try and optimize algae production if you plug in different theoretical nutrient value numbers, a useful tool for the future. This model does not work, but this just gives us an idea what still needs to be improved in our tools.

There have been other studies looking at algae growth models to try and predict biomass for future cultivation markets (James & Boriah, 2010; Huesemann et al. 2013). They have been generally successful in their modelling to match their observed data, and it appears it is because their model is a function of light intensity and temperature. Perhaps if this model could incorporate light intensity and temperature as a larger variable, the model could become a better predictor tool for future use. But there are differences between the model examined in this report and the ones from other reports; the other models are more mechanistic in nature, building off of algae growth factors instead of nutrient growth.

Future studies should include more data sets from other algae production experiments and facilities. Algae Testbed Public-Private Partnership (ATP3) project and Arizona Algae cultivation data sets could be another opportunity to test for future studies, as exploring fits for those projects may give more robust information and conclusion on the model fit capability. Also, attempts to try and improve the model would be a good future project based on the results of this study. Standardized data recording could also come from a well working model that experts can easily understand across all the fields. This would be to quickly and easily use the working model by plugging in the information and records organized previously.

Literary Citations

Beal, C. B., Archibald, I., Huntley, M. E., Greene., C. H., Johnson, Z. I. (2018). Integrating Algae with Bioenergy Carbon Capture and Storage (ABECCS) Increases Sustainability. Earth's Future. DOI: 10.1002/2017EF000704

Beal, C. M., Gerber, L. N., Sills, D. L., Huntley, M. E., Machesky, S. C., Walsh, M. J., Tester, J. W., Archibald, I., Granados, J., Greene, C. H. (2015). Algal biofuel production for fuels and feed in a 100-ha facility: A comprehensive techno-economic analysis and life cycle assessment. Algal Research, 10, 266– 279. https://doi.org/10.1016/j.algal.2015.04.017

Fuss, S., Canadell, J. G., Peters, G. P., Tavoni, M., Andrew, R. M., Ciais, P., et al. (2014). Betting on negative emissions. Nature Climate Change, 4(10), 850– 853. https://doi.org/10.1038/nclimate2392

Greene, C. H., Huntley, M. E., Archibald, I., Gerber, L. N., Sills, D. L., Granados, J., Beal, C. M., Walsh, M. J. (2017). Geoengineering, marine microalgae, and climate stabilization in the 21st century. Earth's Future, 5(3), 278– 284. https://doi.org/10.1002/2016EF000486

Greene, C. H., Huntley, M. E., Archibald, I., Gerber, L. N., Sills, D. L., Granados, J., Tester, J. W., Beal, C. M., Walsh, M. J., Bidigare, R. R., Brown, S. L., Cochlan, W. P., Johnson, Z. I., Lei, X. G., Machesky, S. C., Redalje, D. G., Richardson, R. E., Kiron, V., and Corless, V. (2016). Marine microalgae: Climate, energy, and food security from the sea. Oceanography 29(4). DOI: 10.5670/oceanog.2016.91

Greene, C. H., Monger, B. C., & Huntley, M. E. (2010). Geoengineering: The inescapable truth of getting to 350. Solutions, 1(5), 57– 66.

Hogle, S. L., Dupont, C. L., Hopkinson, B. M., King, A. L., Buck, K. N., Roe, K. L., Stuart, R. K., Allen, A. E., Mann, E. L., Johnson, Z. I., Barbeau, K. A. (2018). Pervasive iron limitation at subsurface chlorophyll maxima of the California Current. Proc Natl Acad Sci U.S.A. : 201813192. DOI: 10.1073/pnas.1813192115

Huesemann, M. H., Van Wagenen, J., Miller, T., Chavis, A., Hobbs, S., & Crowe, B. (2013). A screening model to predict microalgae biomass growth in photobioreactors and raceway ponds. Biotechnology and bioengineering, 110(6), 1583-1594.

Huntley, M. E., Johnson, Z. I., Brown, S. L., Sills, D. L., Gerber, L., Archibald, I., Machesky, S. C., Granados, J., Beal, C. & Greene, C. H. (2015). Demonstrated large-scale production of marine microalgae for fuels and feed. Algal Research. DOI:

10.1016/j.algal.2015.04.016

IPCC. (2013). Climate change 2013: The physical science basis. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.), Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change (p. 1535). Cambridge, UK and New York, NY: Cambridge University Press.

James, S. C., & Boriah, V. (2010). Modeling algae growth in an open-channel raceway. Journal of Computational Biology, 17(7), 895-906.

Karl, T. R., & Trenberth, K. E. (2003). Modern global climate change. Science, 302(5651), 1719– 1723. https://doi.org/10.1126/science.1090228.

Loftus, S. E. and Johnson, Z. I. (2017) Cross-study analysis of factors affecting algae cultivation in recycled medium for biofuel production. Algal Research DOI: 10.1016/j.algal.2017.03.007

Moody, J. W., McGinty, C. M. and Quinn, J. C. (2014). Global evaluation of biofuel potential from microalgae Proc. Natl Acad. Sci. USA 111 8691–6

National Oceanic and Atmospheric Administration (2017a). Global greenhouse gas reference network. United States National Centers for Environmental Information. Retrieved from https://www.esrl.noaa.gov/gmd/ccgg/trends/full.html

National Oceanic and Atmospheric Administration (2017b). Climate at a glance. United States National Centers for Environmental Information. Retrieved from https://www.ncdc.noaa.gov/cag/time-series/global/globe/land/ytd/12/1880-2017

Oreskes, N. (2004). The scientific consensus on climate change. Science, 306(5702), 1686–1686. https://doi.org/10.1126/science.1103618

United Nations (2017), World population prospects: The 2017 revision, key findings and advance tables (Technical Report ESA/P/WP/248), United Nations.

Walsh, B., Ciais, P., Janssens, I. A., Penuelas, J., Riahi, K., Rydzak, F., Vuuran, D. P. V., Obersteiner, M. (2017). Pathways for balancing CO2 emissions and sinks. Nature Communications, 8. https://doi.org/10.1038/ncomms14856

Walsh, M. J., Gerber, Van-Doren, L., Sills, D. L., Archibald ,I., Beal, C. M., Lei, X. G., Huntley. M. E., Johnson, Z. & Greene, C. H. (2016). Algal food and fuel co production can mitigate greenhouse gas emissions while improving land and water-use efficiency. Environmental Research Letters 11:114006.