Biodegradable Dissolved Organic Material

Experiment done by Haley Moon
Presented by Jimmy Tyler
Analysis by Jimmy Tyler, Jake Murphy, and Daniel Smith

Experiment Outline:

Jake: Experiment Background

Daniel: Data Analysis

Jimmy: Applications to Other Experiments

1. Goal is to compare this analysis to others that are analyzing effect of wildfires and anthropogenic factors on nutrient flow through an ecosystem.
2. Goal is to see if we can predict the BDOM of a sample after a treatment using only the properties observable before the treatment.
3. Goal is to generalize our results to other experiments in order to increase applicability.
4. Goal is to motivate further study into specific topics, and address follow-up questions.

Crandall Study Citations:

Citation: Crandall T, Jones E, Greenhalgh M, Frei RJ, Griffin N, Severe E, et al. (2021) Megafire affects stream sediment flux and dissolved organic matter reactivity, but land use dominates nutrient dynamics in semiarid watersheds. PLoS ONE 16(9): e0257733. https://doi.org/10.1371/journal.pone.0257733

Crandall Study Premise:

"Extremely low snowpack, record precipitation, and high average summer temperatures set the scene for the three large wildfires in the Utah Lake watershed that converged into a megafire in 2018." - Crandall, 2021

- The Coal Hollow fire was started on August 4th by a lightning strike, and burned 128 km2
- The Bald Mountain fire was started on August 24th and burned 75 km2
- Pole Creek fire was started on September 6th and burned 413 km2

Together, these wildfires burned a contiguous area of 616 km2, creating a megafire referred to as the Pole Creek Fire Complex.

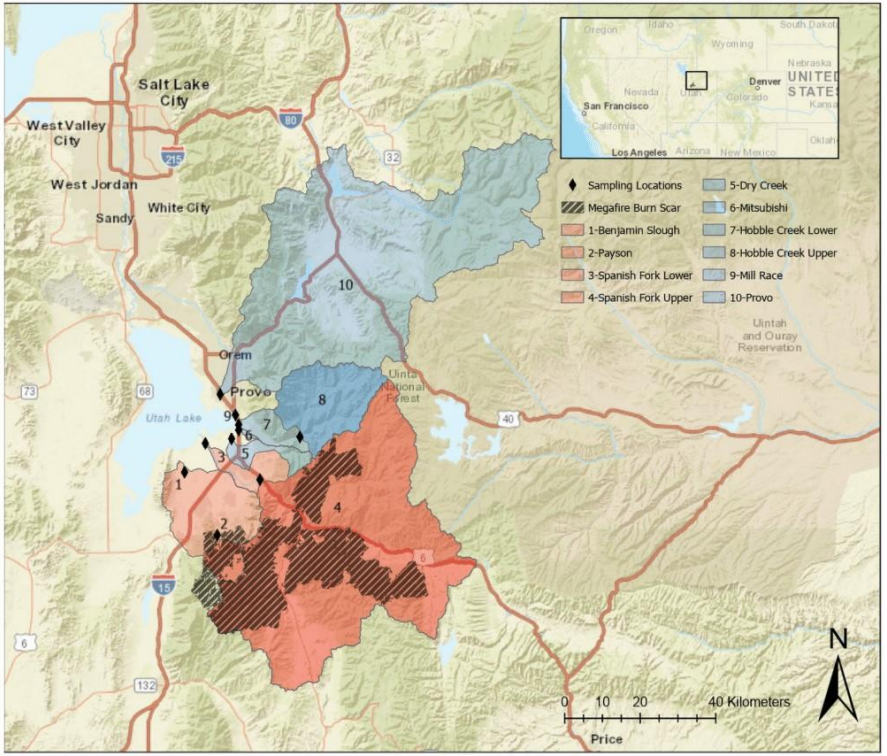


Fig 1. Map showing the megafire burn scar, sampling locations, and contributing watersheds. Watersheds are displayed as either red (burned) or blue (unburned) with numbers indicating the river or stream name in the legend. The map was created using opensource data from the USGS National Map and the Earth Resources Observation and Science Center.

Crandall: Utah Lake “Plume” of Sediment

Crandall Study Stats:

These are the stats for the Crandall Study site-specific data.
A lot of these are similar to the stats that we had for our initial properties that we use in our later machine learning analysis.

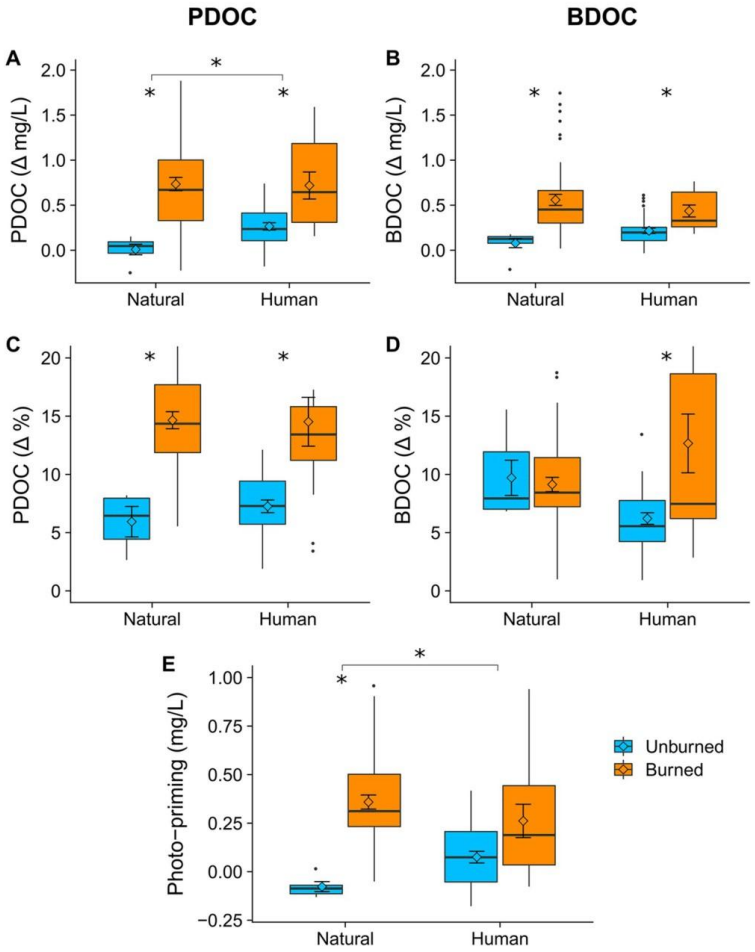
Table 1. Watershed characteristics of sampling sites within the Utah Lake watershed.

Site number and name	Area (km^2)	Burned (%)	Mean slope (°)	Urban (%)	Impervious surface (%)	Agriculture (%)	Forest (%)	Herbaceous (%)
1-Benjamin Slough	319	67	21	10.6	3.7	18.0	38.4	6.4
2-Payson*	54	90	29	2.68	0.0	0.0	82.9	0.5
3-Spanish Fork Lower	1712	24	32	1.69	0.4	8.0	56.3	5.5
4-Spanish Fork Upper*	1650	25	32	0.87	0.1	3.0	57.5	5.3
5-Dry Creek	30	0	2	54.5	23.8	15.0	2.98	8.6
6-Mitsubishi Race	3	0	1	61.2	28.0	0.1	2.81	1.8
7-Hobble Creek Lower	324	0	41	4.48	1.7	1.5	56.7	7.6
8-Hobble Creek Upper*	280	0	42	0.51	0.0	1.2	60	7.2
9-Mill Race	47	0	36	40.6	23.0	0.5	47.1	2.9
10-Provo River	1769	0	27	5.19	0.9	7.0	65.1	3.5

*Sites categorized as "natural" (less than 10% urban and agricultural land use)

Crandall Study Visualization:

- PDOC and BDOC are both measures of carbon content, comparable to our measure of BDOM.
- They are slightly connotatively different, and for anyone curious in this difference, you would have to read the Crandall Study.
- Note the similarity in their findings versus ours, on how fire affects carbon content.



Crandall Study: Results

“Why is burned BDOM so degradable? Contrary to our hypotheses, we observed higher photo and biodegradability at burned sampling locations.” - Crandall, 2021

- Expected that loss of plant matter and the high-temperature pyrogenic compounds during the megafire would result in less reactive compounds in burned watersheds.
- However, pyrogenic organic matter varies widely in composition and reactivity.
- The effects of combustion on organic matter properties depend on initial substrate (e.g. plant or soil sources), the characteristics of the combustion process (e.g. temperature, duration, and percentage consumption), and the degree of processing during subsequent transport to and through the hydrological network.

actual	lower	mid	upper	error	check
5.833333333	2.782855773036862	8.668344389485373	14.35215480426577	2.8350110564853734	Inside
11.31775	2.4694191854450778	9.663555926619003	14.470709067581433	1.6541940733809977	Inside
21.297	2.4724009631363963	9.850681200513332	18.074931354914682	11.446318799486669	Outside
3.0905	2.8627406185821616	3.9328015478523923	8.175140747953648	0.8423015478523923	Inside
4.0655	2.5400461570114268	5.302400032013596	8.11721749736577	1.2369000320135957	Inside
2.599333333	2.781073705242901	3.6665585022979696	7.138369789865203	1.0672251692979695	Outside
5.99	2.8069588791074147	6.37816846596484	7.780660951863208	0.3881684696483945	Inside
1.013	0.9335382022243592	1.754503751988815	1.7741047930406781	0.741503751988815	Inside
3.09925	2.4503756484011112	2.8761883189900592	6.163381339531945	0.22306168100994084	Inside
8.322625	2.806574854547827	8.099064305291407	10.047819960707475	0.22356069470859374	Inside
5.552125	2.806574854547827	5.946747028666241	9.91530234246641	0.39462202866624096	Inside
4.062125	2.7256163762856427	3.6667696194976878	7.04945034874784 95	0.3953553805023122	Inside
3.001	2.6266313032748876	3.327007577126475	9.638009891438294	0.3260075771264752	Inside
4.10875	2.816891459363935	4.100153416996293	7.15494871099766	0.008596583003706 47	Inside
9.364708333	2.80311597225502	10.772969417783 62	14.27480223792678	1.4082610847836214	Inside
1.794166667	1.609030040818547	2.2453919255590273	9.586403582405246	0.4512252585590273	Inside
2.60925	1.7664191921793004	3.056866789217005	9.602982503537703	0.4476167892170051	Inside
5.431125	2.6711435352381954	4.12057226938028	7.066029269880306	1.3105527306197198	Inside
9.22475	2.8069588791074147	8.973250562620555	15.06435174277968	0.25149943737944547	Inside
3.91975	1.80026962003081	2.977848961149271	7.3265556103433 03	0.941901038850729	Inside
7.648375	2.8069588791074147	6.693478104405356	9.481767748716893	0.9548968955946435	Inside
3.788125	2.6266026363045287	4.755016356739382	7.04945034874784 95	0.9668913567393824	Inside
3.697125	2.6371078489111235	5.289601824889967	7.814953184864897	1.592476824889967	Inside
2.046375	2.1180614688820736	4.661395342186634	4.495650733095208	2.6150203421866345	Outside
3.620625	2.61976102226927	3.4289935001046667	7.12747141469159	0.19163149989533323	Inside
8.76018815	2.6593465905027798	8.317747145392735	8.472669950358307	0.44244100460726443	Outside
4.12957865	2.8713643004113822	3.847785681556271	7.445575324297824	0.28179296844372903	Inside
2.825875	2.728382931989415	3.9838976410764078	7.216390855808944	1.1580226410764078	Inside
3.054625	2.63219144928004	3.048701545379097	7.25692405318087 45	0.005923454620903268	Inside
13.41033333	2.8069588791074147	17.581706082568587	22.196897113751085	4.171372752568587	Inside
6.504375	2.8069588791074147	8.24102426732978	12.059711909544122	1.7366492673297813	Inside
11.87954167	2.8709802758517946	11.799498132534952	14.760522677082442	0.08004353746504833	Inside
3.098	2.816891459363935	3.1302328413310527	7.138369789865203	0.032232841331052864	Inside
10.96820833	2.8069588791074147	18.417013074245794	24.941665028416413	7.448804744245795	Inside
7.291166667	2.816891459363935	8.965983459643361	14.430175870209 51	1.6748167926433615	Inside
1.2705	3.3523336349608597	1.2480660154702827	7.062054895661293	0.022433984529717232	Outside
9.08675	2.647365797153615	12.980036549529771	16.815821194264714	3.893286549529771	Inside
5.552125	2.8610476955952744	5.823012292207822	9.837281276522669	0.2708872922078216	Inside
9.22925	2.861431720154862	7.7375709855381 88	10.477221122926839	1.4916790144618126	Inside

Motivation for Machine Learning Analysis:

- Eliminating the time-consuming process of testing and treating water samples.
- Eliminating the treatment wait-time, which can often be quite long. In our study, 28 days.
- Guessing carbon level is the first step in improving BDOM.
- Its cool. Seriously, I had a lot of fun!

This image will be explained further later, but is the output of my machine learning analysis on the test set.

actual	lower	mid	upper	error	check
5.833333333	2.782855773036862	8.668344389485373	14.35215480426577	2.8350110564853734	Inside
11.31775	2.4694191854450778	9.663555926619003	14.470709067581433	1.6541940733809977	Inside
21.297	2.4724009631363963	9.850681200513332	18.074931354914682	11.446318799486669	Outside
3.0905	2.8627406185821616	3.9328015478523923	8.175140747953648	0.8423015478523923	Inside
4.0655	2.5400461570114268	5.302400032013596	8.11721749736577	1.2369000320135957	Inside
2.599333333	2.781073705242901	3.6665585022979696	7.138369789865203	1.0672251692979695	Outside
5.99	2.8069588791074147	6.37816846596484	7.780660951863208	0.3881684696483945	Inside
1.013	0.9335382022243592	1.754503751988815	1.7741047930406781	0.7415037519888815	Inside
3.09925	2.4503756484011112	2.8761883189900592	6.163381339531945	0.22306168100994084	Inside
8.322625	2.806574854547827	8.099064305291407	10.047819960707475	0.22356069470859374	Inside
5.552125	2.806574854547827	5.946747028666241	9.91530234246641	0.39462202866624096	Inside
4.062125	2.7256163762856427	3.6667696194976878	7.04945034874784 95	0.3953553805023122	Inside
3.001	2.6266313032748876	3.327007577126475	9.638009891438294	0.3260075771264752	Inside
4.10875	2.816891459363935	4.100153416996293	7.15494871099766	0.008596583003706 47	Inside
9.364708333	2.80311597225502	10.772969417783 62	14.274802237926 78	1.4082610847836214	Inside
1.794166667	1.609030040818547	2.2453919255590273	9.586403582405246	0.4512252585590273	Inside
2.60925	1.7664191921793004	3.056866789217005	9.602982503537703	0.4476167892170051	Inside
5.431125	2.6711435352381954	4.12057226938028	7.066029269880306	1.3105527306197198	Inside
9.22475	2.8069588791074147	8.973250562620555	15.064351747277968	0.25149943737944547	Inside
3.91975	1.80026962003081	2.977848961149271	7.326555610343303	0.941901038850729	Inside
7.648375	2.8069588791074147	6.693478104405356	9.481767748716893	0.9548968955946435	Inside
3.788125	2.6266026363045287	4.755016356739382	7.04945034874784 95	0.9668913567393824	Inside
3.697125	2.6371078489111235	5.289601824889967	7.814953184864897	1.592476824889967	Inside
2.046375	2.1180614688820736	4.661395342186634	4.495650733095208	2.6150203421866345	Outside
3.620625	2.61976102226927	3.4289935001046667	7.12747141469159	0.19163149989533323	Inside
8.76018815	2.6593465905027798	8.317747145392735	8.472669950358307	0.44244100460726443	Outside
4.12957865	2.8713643004113822	3.847785681556271	7.445575324297824	0.28179296844372903	Inside
2.825875	2.728382931989415	3.9838976410764078	7.216390855808944	1.1580226410764078	Inside
3.054625	2.632191144928004	3.048701545379097	7.25692405 31808745	0.005923454620903268	Inside
13.41033333	2.8069588791074147	17.581706082568587	22.196897113751085	4.171372752568587	Inside
6.504375	2.8069588791074147	8.24102426732978	12.059711909544122	1.7366492673297813	Inside
11.87954167	2.8709802758517946	11.799498132534952	14.760522677082442	0.08004353746504833	Inside
3.098	2.816891459363935	3.1302328413310527	7.138369789865203	0.03223284133105 2864	Inside
10.96820833	2.8069588791074147	18.417013074245794	24.941665028416413	7.448804744245795	Inside
7.291166667	2.816891459363935	8.965983459643361	14.430175870209 51	1.6748167926433615	Inside
1.2705	3.3523336349608597	1.2480660154702827	7.062054895661293	0.022433984529717232	Outside
9.08675	2.647365797153615	12.980036549529771	16.815821194264714	3.893286549529771	Inside
5.552125	2.8610476955952744	5.823012292207822	9.837281276522669	0.2708872922078216	Inside
9.22925	2.861431720154862	7.7375709855 38188	10.477221122926839	1.4916790144618126	Inside

The only explanatory variables we considered were initial measurements. Though some variables, like site-specific data didn't change, others did, such as optical properties of samples. These were held in stasis using acidic cultures, and measured at the same time as other variables.

Explanatory Variables:

- Season(Trial), Treatment(CL,CD,NL,ND)
- Optical Properties: Fluorescence Index, BIX, HIX, Absorbance at 254 nanometers
- Average Precipitation over the last 12 months
- Site Specific Data: Drainage Area, Forestation Percentage, Developed Land Percentage, Land Percentage Impervious to fire, Herbaceous Land Percentage, slope
- Burn Level
- Sample Specific Data: Carbon Level at initial measurement (t0)

Response Variables:

- Carbon Level after treatment (t28)
- Percentage of change in carbon level (BDOM)

actual	lower	mid	upper	error	check
5.833333333	2.782855773036862	8.668344389485373	14.35215480426577	2.8350110564853734	Inside
11.31775	2.4694191854450778	9.663555926619003	14.470709067581433	1.6541940733809977	Inside
21.297	2.4724009631363963	9.850681200513332	18.074931354914682	11.446318799486669	Outside
3.0905	2.8627406185821616	3.9328015478523923	8.175140747953648	0.8423015478523923	Inside
4.0655	2.5400461570114268	5.302400032013596	8.11721749736577	1.2369000320135957	Inside
2.599333333	2.781073705242901	3.6665585022979696	7.138369789865203	1.0672251692979695	Outside
5.99	2.8069588791074147	6.37816846596484	7.780660951863208	0.3881684696483945	Inside
1.013	0.9335382022243592	1.754503751988815	1.7741047930406781	0.741503751988815	Inside
3.09925	2.4503756484011112	2.8761883189900592	6.163381339531945	0.22306168100994084	Inside
8.322625	2.806574854547827	8.099064305291407	10.047819960707475	0.22356069470859374	Inside
5.552125	2.806574854547827	5.946747028666241	9.91530234246641	0.3946220286624096	Inside
4.062125	2.7256163762856427	3.6667696194976878	7.04945034874784 95	0.3953553805023122	Inside
3.001	2.6266313032748876	3.327007577126475	9.638009891438294	0.3260075771264752	Inside
4.10875	2.816891459363935	4.100153416996293	7.15494871099766	0.008596583003706 47	Inside
9.364708333	2.80311597225502	10.772969417783 62	14.27480223792678	1.4082610847836214	Inside
1.794166667	1.609030040818547	2.2453919255590273	9.586403582405246	0.4512252585590273	Inside
2.60925	1.7664191921793004	3.056866789217005	9.602982503537703	0.4476167892170051	Inside
5.431125	2.6711435352381954	4.12057226938028	7.066029269880306	1.3105527306197198	Inside
9.22475	2.8069588791074147	8.973250562620555	15.064351747277968	0.25149943737944547	Inside
3.91975	1.80026962003081	2.977848961149271	7.3265556103433003	0.941901038850729	Inside
7.648375	2.8069588791074147	6.693478104405356	9.481767748716893	0.9548968955946435	Inside
3.788125	2.6266026363045287	4.755016356739382	7.04945034874784 95	0.9668913567393824	Inside
3.697125	2.6371078489111235	5.289601824889967	7.814953184864897	1.592476824889967	Inside
2.046375	2.1180614688820736	4.661395342186634	4.495650733095208	2.6150203421866345	Outside
3.620625	2.61976102226927	3.4289935001046667	7.127471414691559	0.19163149989533323	Inside
8.76018815	2.6593465905027798	8.317747145392735	8.472669950358307	0.44244100460726443	Outside
4.12957865	2.8713643004113822	3.847785681556271	7.445575324297824	0.28179296844372903	Inside
2.825875	2.728382931989415	3.9838976410764078	7.216390855808944	1.1580226410764078	Inside
3.054625	2.632191444928004	3.048701545379097	7.25692405318087 45	0.005923454620903268	Inside
13.41033333	2.8069588791074147	17.581706082568587	22.196897113751085	4.171372752568587	Inside
6.504375	2.8069588791074147	8.24102426732978	12.059711909544122	1.7366492673297813	Inside
11.87954167	2.8709802758517946	11.799498132534952	14.760522677082442	0.08004353746504833	Inside
3.098	2.816891459363935	3.1302328413310527	7.138369789865203	0.032232841331052864	Inside
10.96820833	2.8069588791074147	18.417013074245794	24.941665028416413	7.448804744245795	Inside
7.291166667	2.816891459363935	8.965983459643361	14.43017587020951	1.6748167926433615	Inside
1.2705	3.3523336349608597	1.2480660154702827	7.062054895661293	0.022433984529717232	Outside
9.08675	2.647365797153615	12.980036549529771	16.815821194264714	3.893286549529771	Inside
5.552125	2.8610476955952744	5.823012292207822	9.837281276522669	0.2708872922078216	Inside
9.22925	2.861431720154862	7.7375709855538188	10.477221122926839	1.4916790144618126	Inside

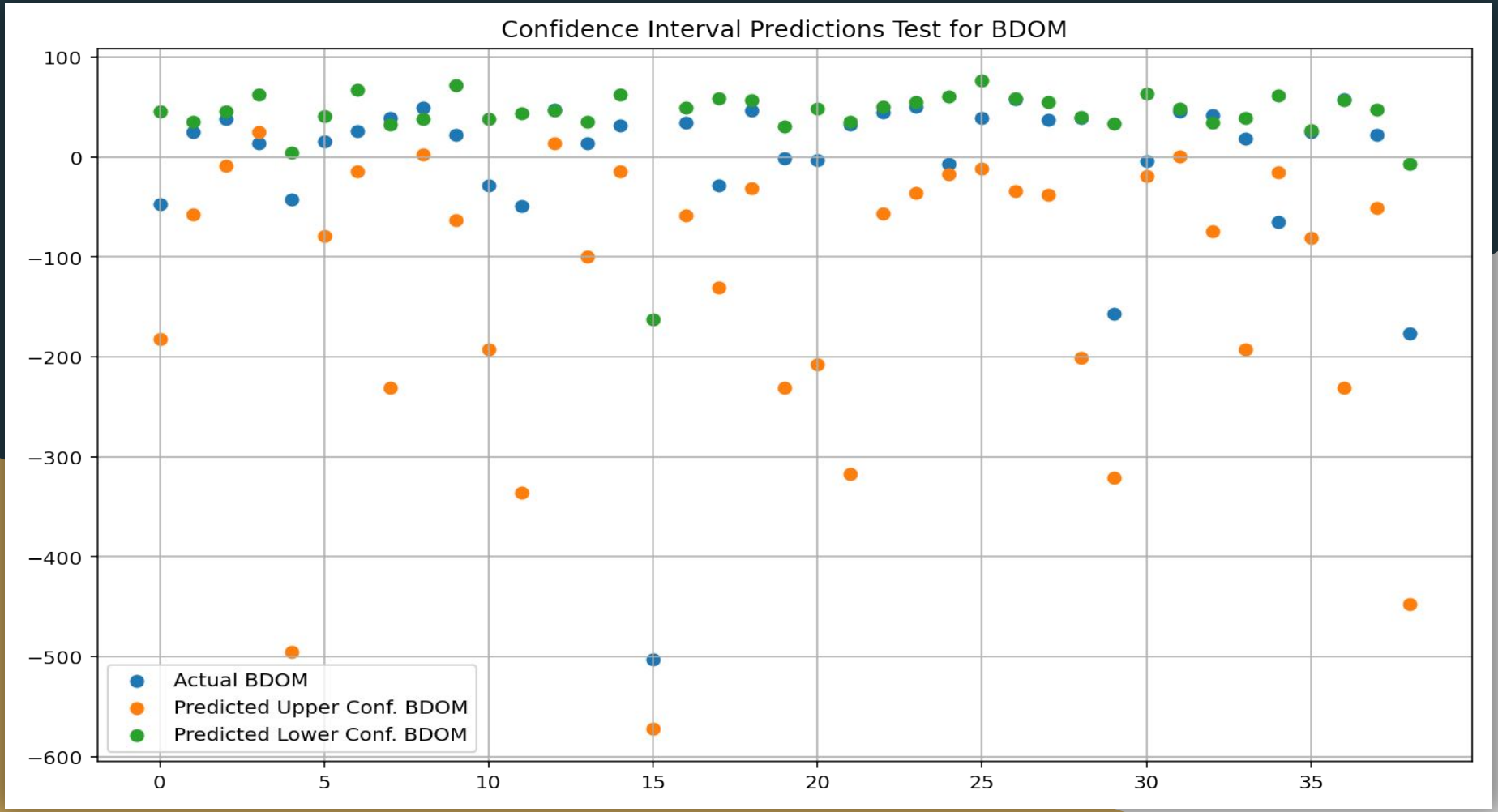
Methods: Main Package utilized was Sci-Kit Learn

Data was Preprocessed:

- First, our goal was to predict t28 and BDOM
- Second, we shuffled the samples to ensure no confounding pattern in the data would skew our results.
- Third, encoding to categorically represent our treatment values
- Fourth, we split our data into four groups at a rate of 80/20.
- Finally, we utilized standard feature scaling of numerical values for increased accuracy of our machine learning algorithm.

Data was treated using Multivariable Linear Machine Learning Techniques:

- Sci-Kit Learn Linear Regression algorithms were implemented
- Further tests were conducted using dnn's in an unsupervised environment from Sci-Kit Learn, but these were not as accurate as the Linear methods.
- Potential for dnn's are still possible, discussed later.


Confidence Interval Predictions Test for BDOM
100
0
−100
−200
−300
−400
−500
−600
0
5
10
15
20
25
30
35
Actual BDOM
Predicted Upper Conf. BDOM
Predicted Lower Conf. BDOM

Applications to past studies

- Historically forest fires cleared old growth, but with Climate Change, worry is that changing fires will affect nutrient flow through an environment
- As fires change, nutrient flow through the system changes, and hence recovery from fires, and regrowth, changes
- One major concern is that nutrients will flow out of a system based off of water-shed, as opposed to stay in the system. Crandall Study suggests BDOM is a measure of this
- What the Crandall study didn't consider which we studied were a few additional factors:
    - The burn level of a specific location
    - The affect changing over time
    - The treatments Haley conducted on the water samples
    - Correlation between specific optical properties of water samples and their BDOM

Motivations for Future Studies

- One extremely interesting problem discussed in the Crandall study is the changing shape of modern classified “megafires”.
- Megafires typically would be not very elliptical in shape, but rather would “branch out”. Speculation is made in the paper that modern megafires are now becoming more elliptical in nature due to the affects of man-made agricultural and urban development.
- The reason why this might be a problem, is that regrowth may be a factor of distance vegetation seed has to travel to get to a fertile soil, as well as how much of the BDOM in the water is retained in that time period.
- The question then remains to determine if the convexity of a fire pattern is correlated to its BDOM drop over time.

Motivations for Future Studies

- We could not determine causality between topological data and BDOM nutrient flow. This is because we didn’t do experiments controlling for other variables.
- This gap is an important one to try and bridge if our data and models are going to have any widespread application.
- If we could predict the flow of nutrients through a system before that has happened, we can begin to take steps to intervene where necessary.
- However, experimental studies, as opposed to observational studies, are in this area of course unethical, as it would constitute starting megafires.

Motivations for Future Studies

Future Possible Improvement for the Deep Neural Network:

- If we utilized a mixture of supervised/unsupervised learning in a deep neural network with dedicated layers its possible that this could keep the upper bound on the t28 value from getting so high. This might eliminate the skew
- Our data was unbalanced. This could cause a bias in the estimate of the carbon level, as the underlying dnn simply tries to optimize the loss function by any means necessary
- Its highly likely more data would allow for narrowing the confidence interval we found to a more applicable level

Questions/Comments?

Thank you very much!