Oliver Evans
Dr. Shane Rogers
Clarkson REU 2016
Weekly Progress Report
June 16, 2016

1   Research Goals

I am working with Dr. Shane Rogers to create a mathematical model of how light affects the growth and composition of the sugar kelp, S. latissima, as part of a larger project studying the feasibility of utilizing macroalgae as a productive nutrient sink for wastewater effluent in ocean outfalls. We are also investigating the potential economic and environmental impacts of using this kelp to create biofuel, high-value chemical products, food, and livestock feed among other uses.

I will begin by studying the attenuation of light resulting from dissolved particulate matter in the water, as well as the absorption of light by the kelp plants. Once this model is formulated using experimental data collected at SINTEF in Trondheim, Norway as well as elsewhere in literature, we will use our findings to inform the overall model of kelp growth.

2   Current Progress: Experimental Data Analysis

2.1   Experimental Setup

Thus far, besides beginning to study literature on the subject, I have been analyzing data collected from an experiment that Dr. Rogers and his colleagues performed several weeks ago in the sea near Trondheim to measure attenuation with and without the influence of kelp shading. The experimental setup was as follows. A rope is attached to a buoy and allowed to hang freely in the water. Fixed to the rope are light sensors spaced 1m apart from one another from 1m below the water's surface to 8m. Samples are collected simultaneously from all sensors every 10 seconds for several hours. A second such rope is placed nearby in the water during the same time, but with one kelp plant attached to it. Similar measurements are taken. After a certain time, a second kelp plant is attached to the rope. Similar measurements are taken once again.

2.2   Data and Model

I was given raw data from the sensors separated into one text file for each light sensor, each containing many lines with the time and intensity of the individual measurements. I have used Python to perform some analysis on the data collected during the previously described experiment, and generate plots to visualize the data. I divided the raw data into 3 datasets:

- **Control**: first rope without kelp,

- **1 Kelp**: second rope with one kelp plant

- **2 Kelp**: second rope with two kelp plants

Note that these datasets are not the same length. The sensors on the first rope were collecting data for a significantly longer period of time which encompassed the time periods when the other two datasets were being collected, as shown in the time plots below.

For each dataset, we sought to calculate the light attenuation according to the model:

$$I(z) = I_0 e^{-kz} \tag{1}$$

where I_0 is the intensity of light on the surface of the water, and k is the attenuation coefficient.

2.3 Regression Technique

By doing exponential regression, these two parameters, I_0 and k can be calculated. Notice that light is not actually being collected at the surface of the water, rather it is inferred based on the measurements at other depths.

These parameters can be calculated at each point in time in each dataset. This is done by converting the exponential model to a linear model by taking the logarithm of equation 1, which gives

$$\ln I(z) = \ln I_0 - kz \tag{2}$$

Then, a simple least squares fit can be calculated by minimizing the sum, $R(k, I_0)$ of the squared differences between our linear model and the logarithm of the data points, given by the following function of our fit parameters.

$$R(k, I_0) = \sum_{n=1}^{N} (\ln I_n - \ln I_0 + kz)^2 \tag{3}$$

where z_n are the depths at which light is collected (i.e. 1m - 8m), I_n are the measured intensities, and N is the number of depths (i.e. 8). Let $R(k, I_0)$ be called the residual for the time step at we are performing the fitting. By taking partial derivatives with respect to both fit parameters, the residual can be minimized, producing the ideal fit (according to this definition of the residuals). Setting both derivatives to 0, we get the following matrix system of equations.

$$\begin{bmatrix} \sum_{n=1}^{N} z_n^2 & \sum_{n=1}^{N} z_n \\ \sum_{n=1}^{N} z_n & N \end{bmatrix} \begin{bmatrix} k \\ I_0 \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N} z_n \ln I_n \\ \sum_{n=1}^{N} \ln I_n \end{bmatrix} \tag{4}$$

2.4 Statistical Analysis

For each dataset, the previously described regression can be performed for each time step. This gives as many values of the fit parameters for the dataset as there are time steps. The mean (μ) and standard deviation (σ) are given for both parameters for each dataset. Statistical description is also given of the R^2 values and of the raw intensity measurements (at all depths).

	k	I_0	Residuals	Intensity
μ	0.27	3.6+04	2.2e+09	1e+04
σ	0.046	2.7e+04	4.5e+09	1.3e+04

Figure 1: Statistical info for **Control**

	k	I_0	Residuals	Intensity
μ	0.34	5.6e+03	1.5e+09	2.2e+03
σ	0.25	2.1e+04	6.7e+09	1.3e+04

Figure 2: Statistical info for **1 Kelp**

	k	I_0	Residuals	Intensity
μ	-0.65	3.4	2.9e+05	95
σ	0.12	4.3	5.2e+05	1.7e+02

Figure 3: Statistical info for **2 Kelp**

2.5 Visualization

Several different visualization techniques are implemented below for the experimental data and for the fitted parameters

2.5.1 Raw Data

Light intensity is plotted over time for each of the 3 datasets. Each sensor is represented as a separate line.
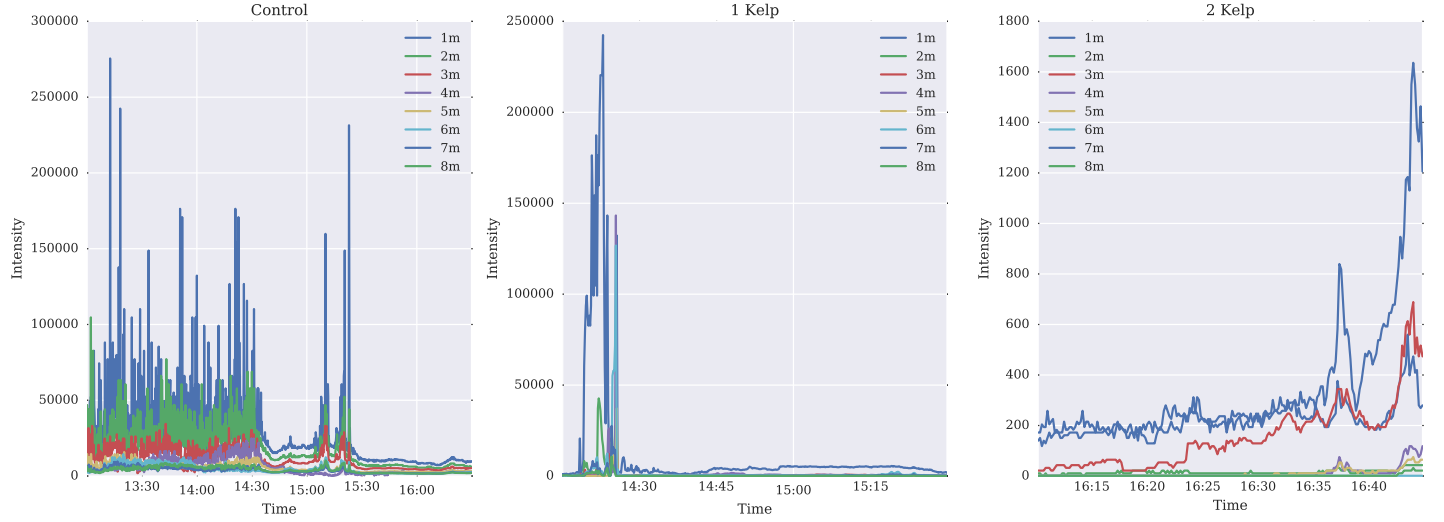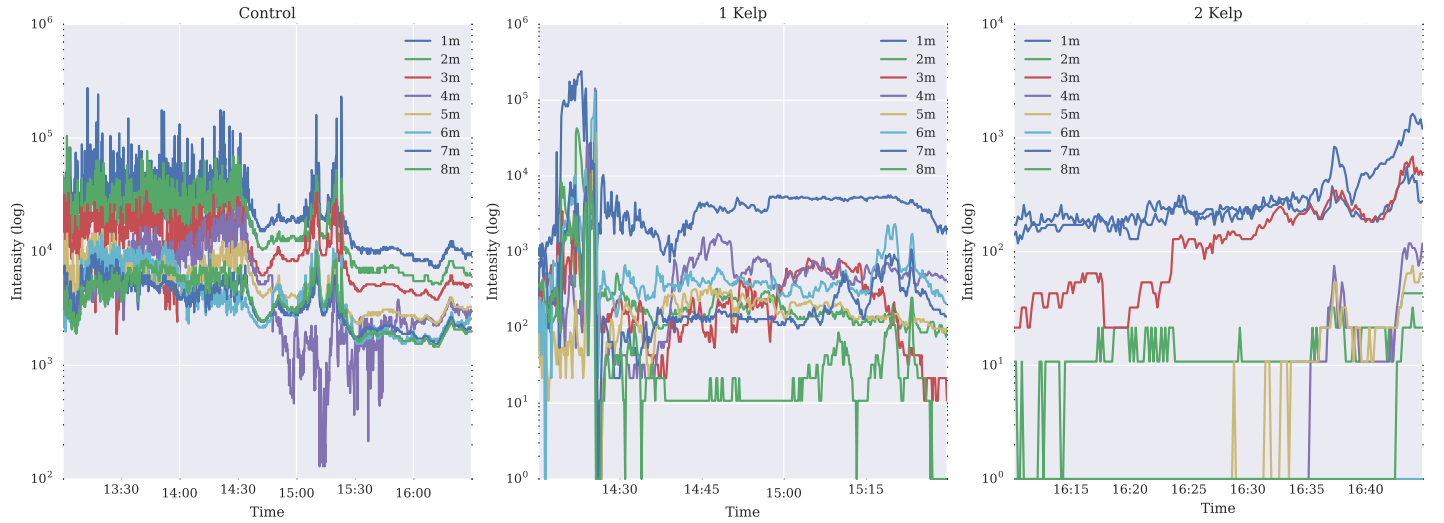

Figure 4: Raw data: regular scale

Figure 5: Raw data: log scale

2.5.2 Parameters over time

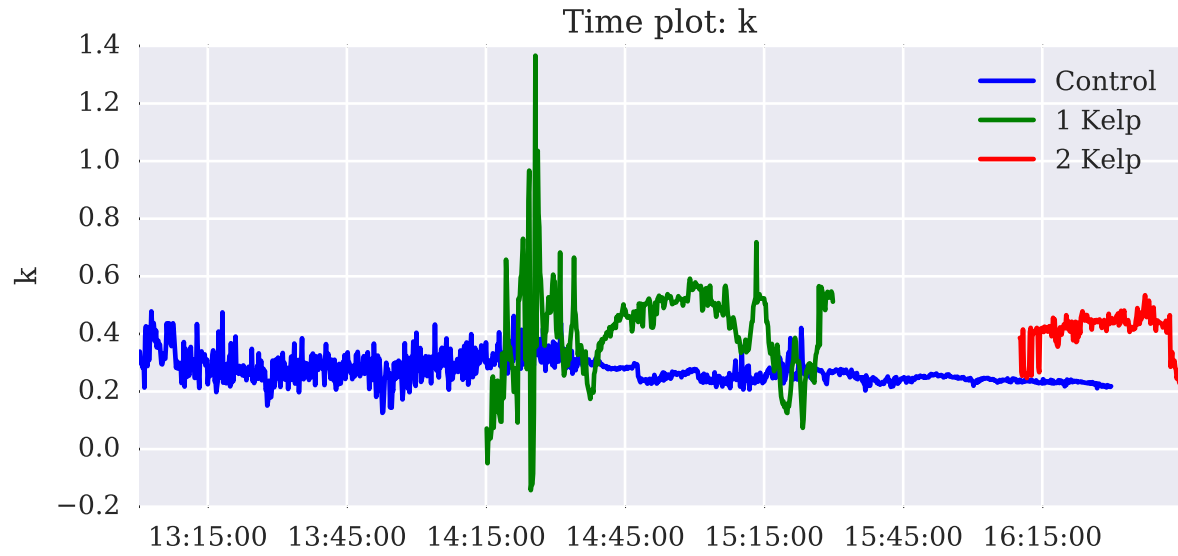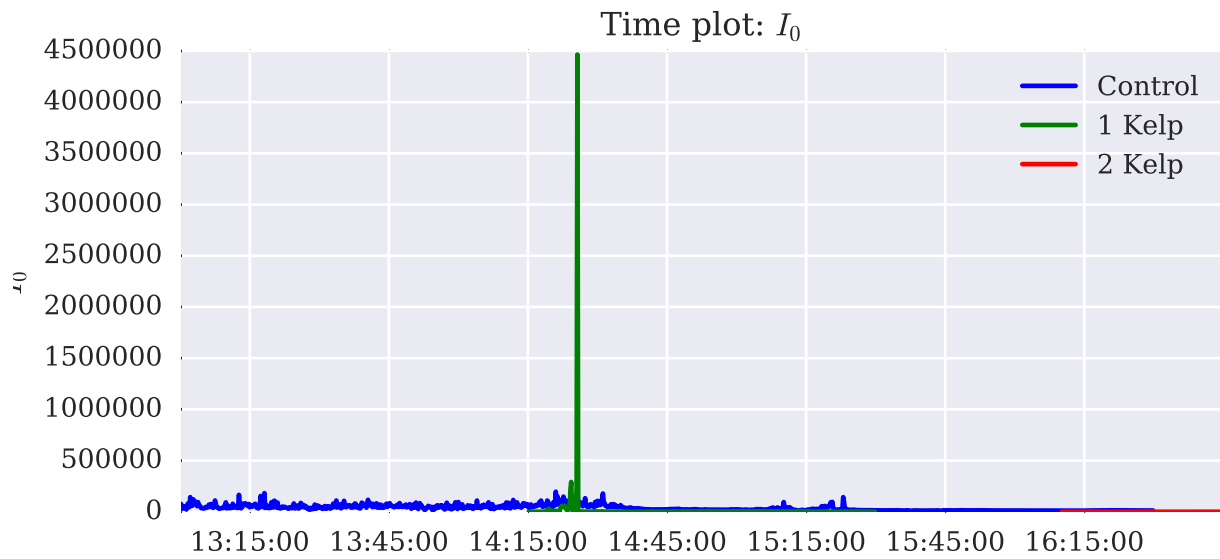The fitted parameters are plotted over time. R^2 values are also plotted.

Figure 6: k over time
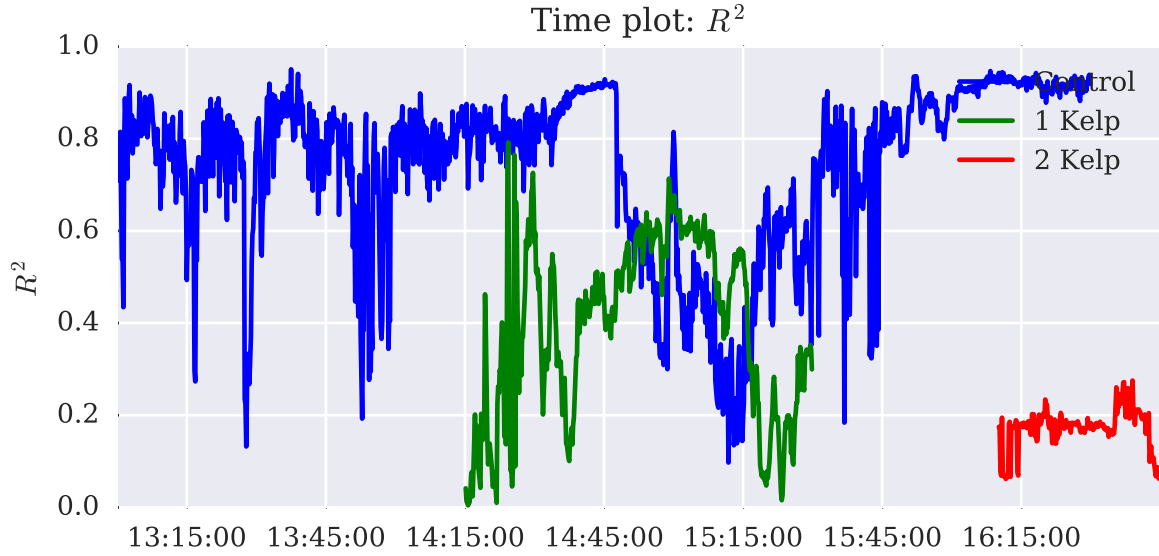

Figure 7: I_0 over time

Figure 8: R^2 over time

2.5.3 Attenuation Coefficient Distribution

The distributions of k is shown using both bar plots and kernel density estimates (KDEs), in which each data point is represented by an exponential function of width similar to the bin width of the histogram, and the normalized sum of the exponentials is drawn to continuously represent the distribution.
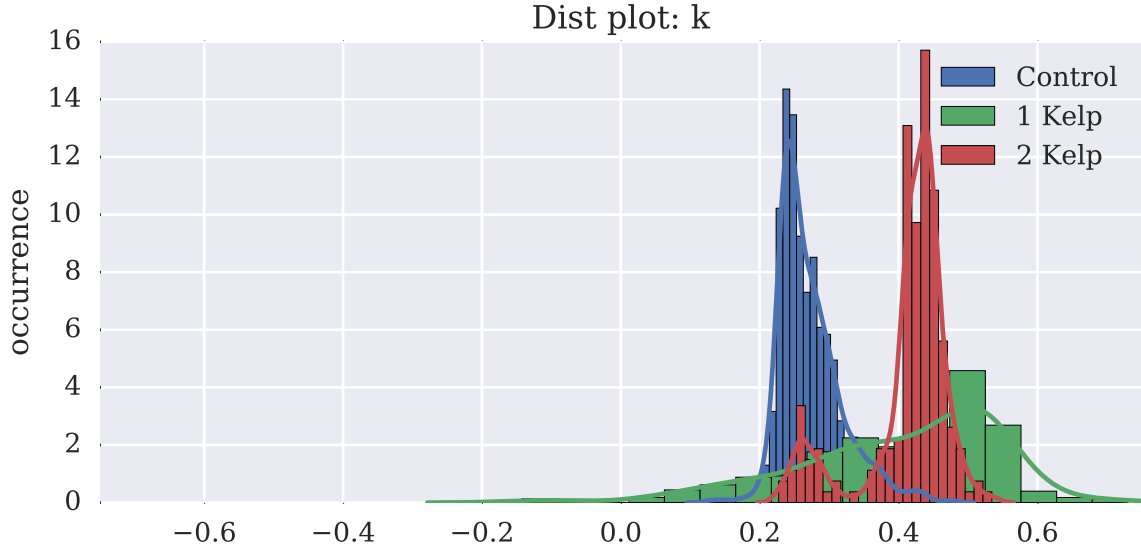

Figure 9: Distribution of fitted values for attenuation coefficient

2.5.4 Joint Plots

The distributions of k and I_0 are plotted together. In the first following figure, all three datasets are plotted together. In the following three plots, k and I_0 are plotted using individual (1d) KDEs on the side and top, and a joint (2D) KDE represented by a contour plot in the center.
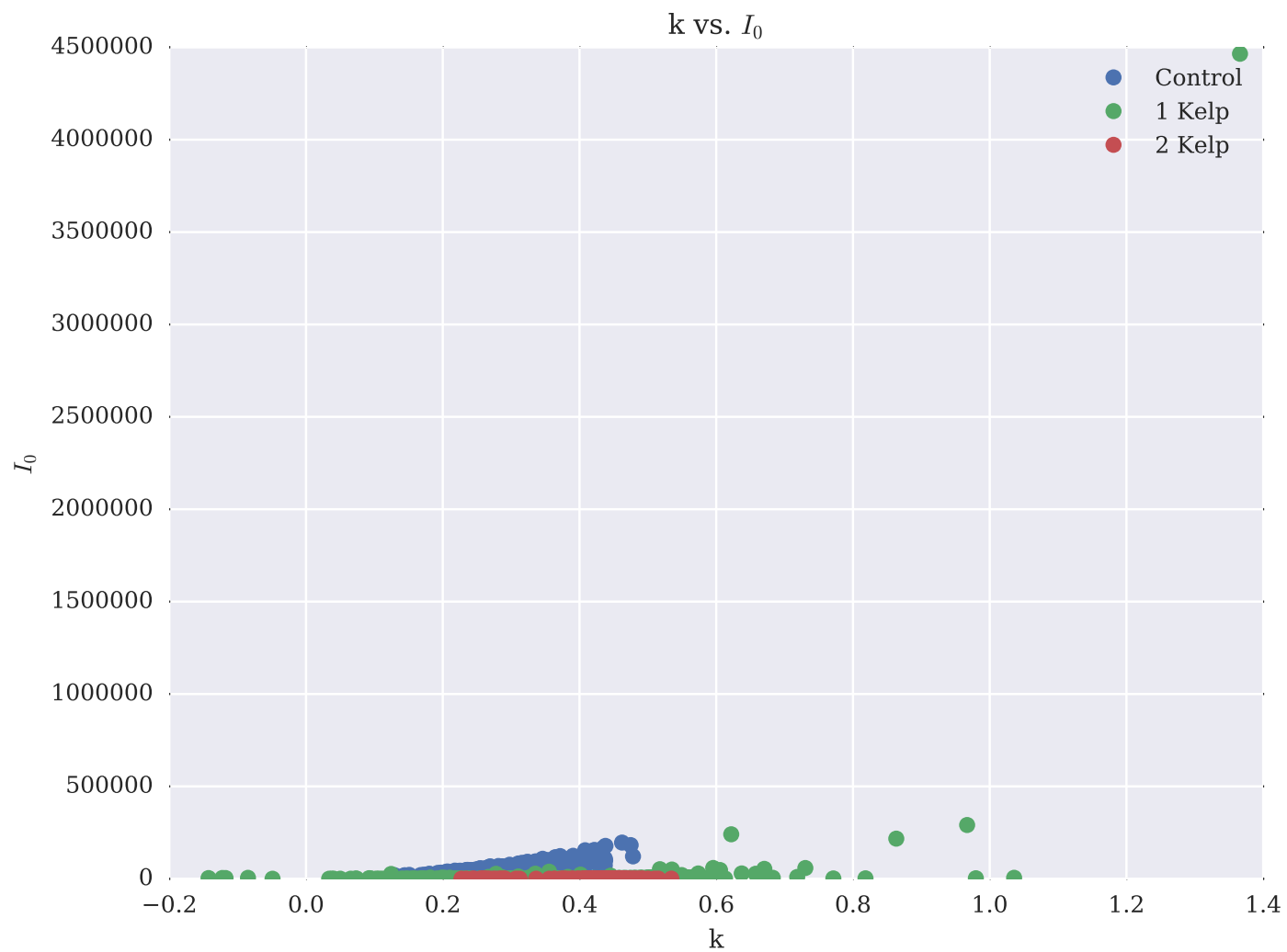
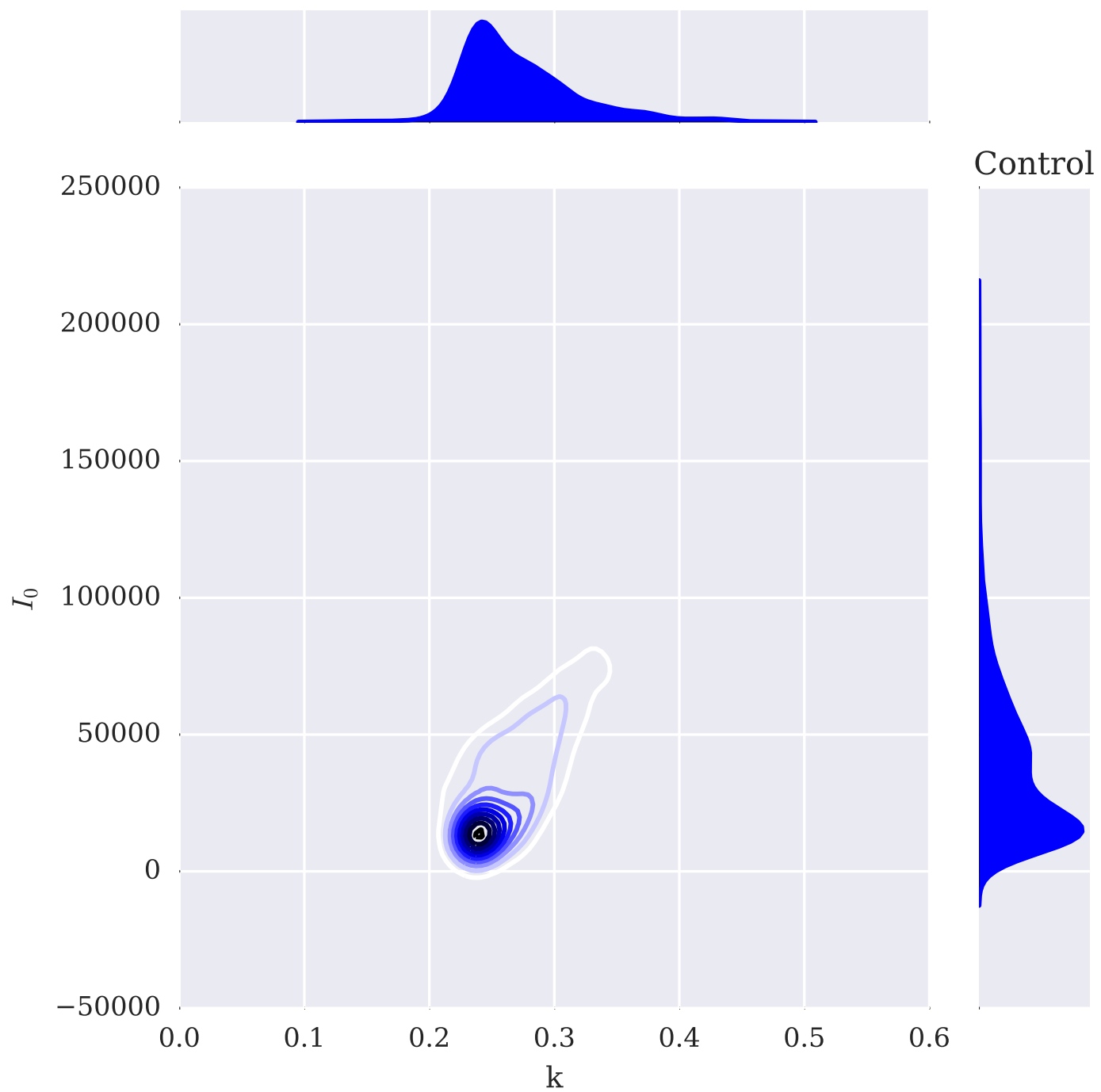Figure 10: Joint distribution of k and I_0 for all datasets

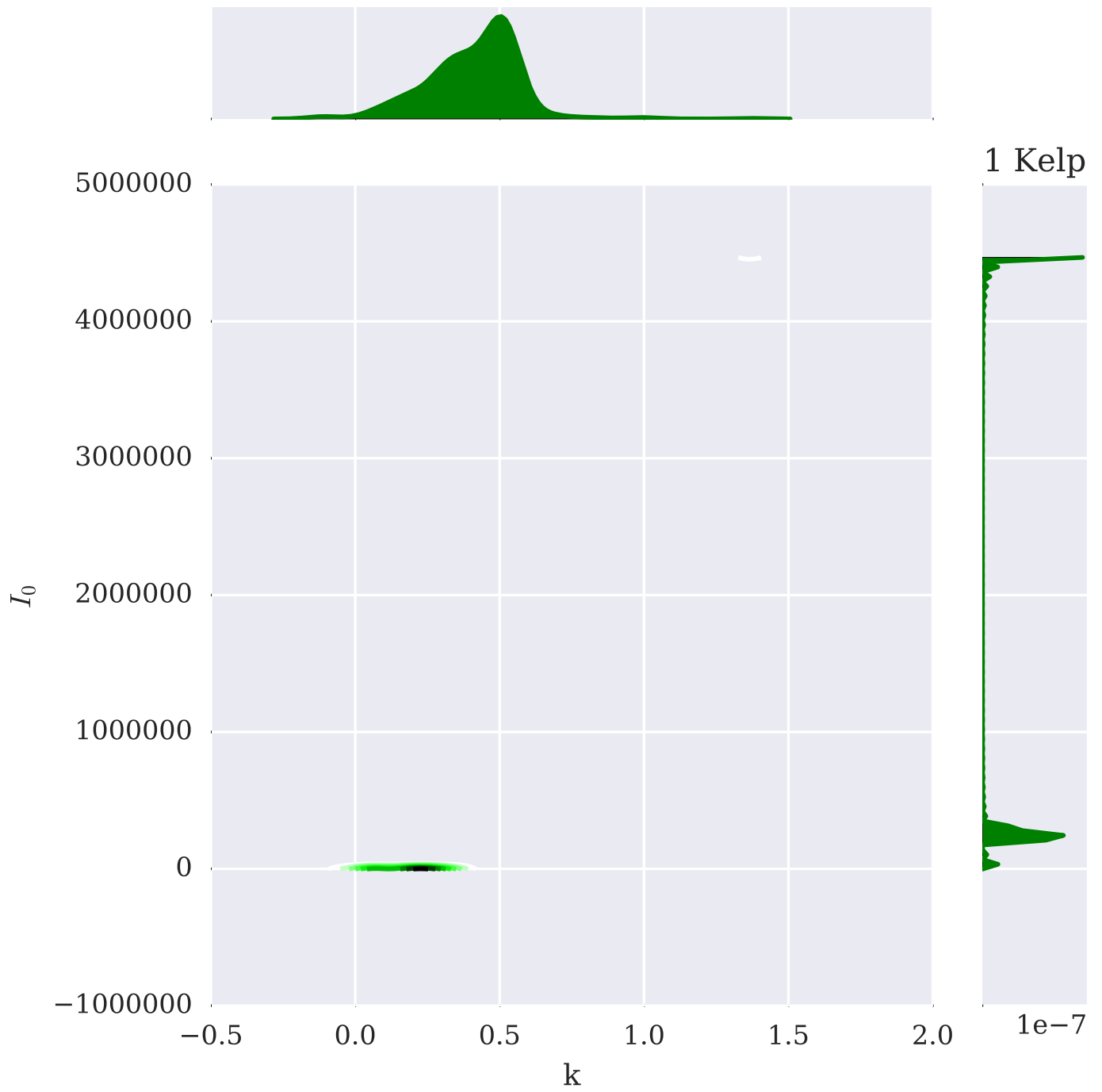Figure 11: Joint distribution of k and I_0 for **Control**

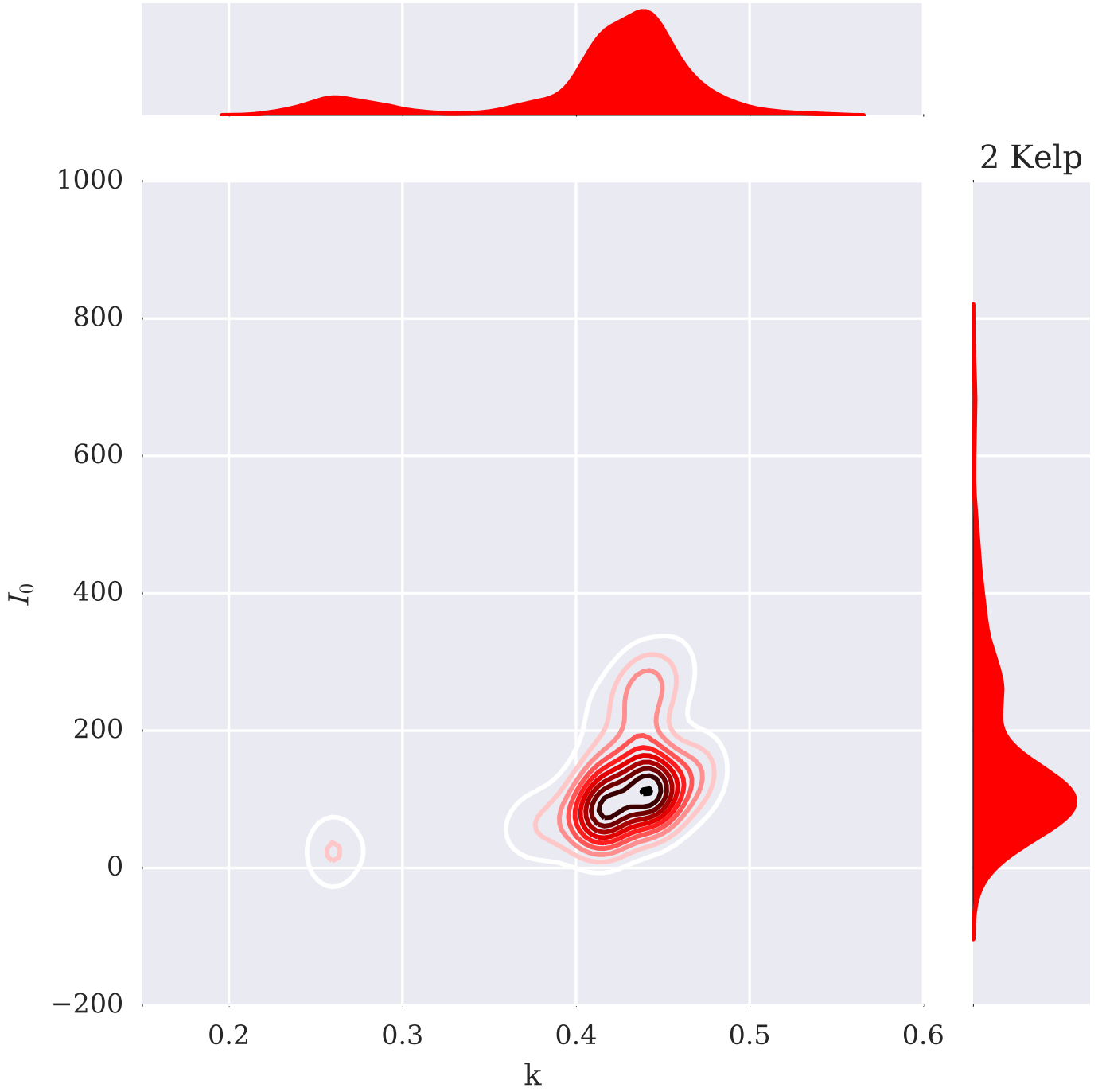Figure 12: Joint distribution of k and I_0 for **1 Kelp**

Figure 13: Joint distribution of k and I_0 for **2 Kelp**

2.6 Interpretation

As is evident from the plots, all of the data is very noisy. The data does not fit the model perfectly, although in some cases the model is fairly successful. In particular, the attenuation measured by the sensors on the rope without kelp fits the model quite well, as can be seen by the relatively constant value for k in figure 6, as well as by the small standard deviation shown in figure 1. The model was fairly successful for the rope with 1 kelp plant. The data is certainly noisier than the control data. In figure 13, it can be seen that the joint KDE for k and I_0 seems to have 2 local maxima, which is symptomatic of a lack of consistency in the fitting. For the rope with 2 kelp plants, though, the analysis did not produce clear results. As is shown in the figures and the statistical table, negative values of k are calculated. This would imply that light intensity is increasing with depth. Furthermore, the raw intensities as well as the calculated values for I_0 are several orders of magnitude lower than in the other cases. It is not clear why this is the case. Some possible cases are:

- Kelp density varies with depth, so perhaps there is more kelp towards the surface than there is near the bottom of the rope.

- We are using the wrong time frame for this dataset. Perhaps the experiment with 2 kelp plants actually occurred at a different point during the day than the section being considered.

- The rope was placed in the water upside down. This could explain both the increasing intensity and lower magnitudes.

- Some unknown event was disrupting the experiment, e.g. large waves, fish, other boats, etc.

We will repeat the experiment on June 11th. Similar analysis will be performed with the new data, and the results will be compared.

3 Thanks for reading!

I'm very excited to be working on this project this summer, and looking forward to a very interesting and productive several weeks!