Title: To what extent does proximity to the Clifton Pier power station positively correlate with higher percentage algae coverage?

Research Question

To what extent does proximity to the Clifton Pier power station positively correlate with higher percentage algae coverage?

Introduction

Coral reefs are under threat as a habitat: a study by the International Union for the Conservation of Nature of 88 sites in the Caribbean showed 34.8% live coral cover in 1970, 19.1% in 1984 and 16.3% in 2011 (Jackson, Donovan, Cramer & Lam 2014)

Algae need nutrients and sunlight to grow. Fertiliser contains nutrients (nitrates and phosphates) used on farms, which collect into the water table of any area. Where the Mississippi meets the Gulf of Mexico, for example the loss of momentum allows all these nutrients to collect and sit at the surface: combined with sunlight, this is sufficient to cause algal blooms. (DUJS) These are defined as "extremely high cell densities (20 000 to 100 000 cells per millilitre) [of Algae], where the proliferation of algae is dominated by a single species..."(Graham, 2010) Effectively, this is a large visible algae collection. This pattern, where collected nutrients lead to algal bloom is called eutrophication ("natural or artificial addition of nutrients to bodies of water and to the effects of the added nutrients") (USGS 2014).

Qualitatively, a combination of these trends can be seen in the southwest corner of New Providence, Bahamas. Algae growth is extensive throughout the marine environment, specifically in the vicinity of the Clifton Pier power plant and Clifton Heritage Park. Of note in the surrounding environment is a Bunker C leak from Clifton Pier power plant, the tongue of the ocean (a 2 kilometre deep submarine canyon). According to the Birkeland, Reimer and Young (1976) study by the EPA, there is a positive correlation between the presence of Bunker C oil and algae growth: the investigation aims to quantify what, if any, impact the oil leak has on algal cover of the reef.

The purpose of this investigation is to determine to what extent proximity to the Clifton Pier power station positively correlates with higher percentage algae coverage? The oil leak at Clifton Pier is composed, at least partially, of Bunker C. This has been confirmed by oil found underwater rather than on the surface, which, according to the EPA (NOAA NO 6 Fuel Oil 2016) is a property displayed by Bunker C. The EPA study presents three possible explanations for this correlation: one is that the viscosity of the Bunker C (due to its long hydrocarbon chain) allows algae to maintain a hold on surfaces, the second is that something in the hydrocarbon chain kills off reef biomass, which leads to an excess of nutrients in the biome, or the third is that hydrocarbons contains the requisite nutrients themselves. (Birkeland, Reimer, Young, 1976)

Algal coverage is inversely correlated with coral health, so elevated algal coverage bodes poorly for coral and by extension reef health. (NEPA, 2014) Coral reefs house 25% percent of all marine species, (NOAA Biodiversity, 2015) yet only comprise about 0.01% of the ocean. Given the extent of diversity they have, a threat to them is a threat to marine biodiversity, which sustains fishing and tourism around the world. This observed trend, towards increased algal and decreased coral cover, poses a threat to tourism, the chief industry of The Bahamas. The Bahamas was the destination of 5.8 million visitors in 2012 (Hartnell, 2012) Most tourists snorkel, dive or visit the beach, hoping to see reef life of the sort that doesn't exist in their own regions. Therefore, any significant threat to marine habitats, health-wise or aesthetically, is a direct threat to the economic security of Bahamas. Additionally, as someone who learned to snorkel and dive in Clifton bay, it is hard to see such inaction on something so clearly destructive. Anecdotally, over the past 7 years in Clifton bay, there has been a marked increase in algal coverage coupled with increased dead coral.

Scientifically quantifying percentage coverage is typically done using a transect line and quadrats. As this investigation aims to quantify algal blooms on reefs of varying proximity to the oil spill, percentage cover will be compared by use of photo-quadrats, as it is less time consuming underwater.

Hypothesis:

Statistical analysis should show no correlation between proximity to the plant and algae cover.

Studies suggest that a positive correlation exists between contact with bunker c and algal growth. Greater proximity to the plant means more frequent contact with bunker c: in theory, if the oil holds the excess nutrients necessary for algae growth, the more oil there is, the more the algae should grow. Therefore the alternate hypotheses are either that algae growth is positively correlated to proximity to Clifton pier (the "point source pollution"), which would reaffirm the Birkeland Reimer and Young (1976) data, or a second alternate hypothesis, is that proximity to the plant could be inversely correlated to algal coverage. This is not supported by the Birkeland et Al. study, though it is possible.

Materials

- 1x 0.25 square metre transect (.5 x .5 metres)
- 1x Nikon Aw120
- 1x SD Card (>4 gb)
- 1x white slate
- 1x 20 metre transect line
- 1 kilo excess weight (to weigh down transect)
- Random number generator
- Microsoft Excel software
- 2 Divers
- Dive gear
  - 2x BCD
  - 2x Regulator
  - 2x Required weights
  - 2x Dive computers
  - 2x Wetsuits
  - 2x Masks/Fins
  - Underwater compass

Safety, Ethical and Environmental Considerations

Ensure following Predive check is performed with buddy before entering water

- BCD
  - Check that your BC inflates and deflates correctly in order to ensure proper buoyancy control for divers
- Weights
  - Check that correct weights are present and sufficient to sink the diver.
- Releases
  - Check that all releases (stomach, chest, etc..) are functional so that in the event of an emergency, the diver can quickly ascend.
- Air
  - Check that air is turned on and regulator is working.

At no time should an insufficiently trained diver dive without a buddy or in conditions above what they have been certified to manage. This poses obvious safety risks, but also ethical considerations arise, those of endangering a rescue (or recovery) team, and environmental hazards to marine life and ecosystems of a struggling diver.

The experiment carries with it the safety risks associated with fieldwork (specifically diving). Environmentally, the methodology is minimally invasive: nothing will be removed or added to the ecosystem, and care should be taken to prevent impact with corals or other life.

Methods

1. 6 random x values were generated using an online tool
2. Divers put on appropriate gear (masks, regulators, etc), conducted a safety check (as described in safety considerations above)
3. Divers entered water carrying quadrats and transect lines, along with the random x values on a slate.
4. Divers navigated to survey site using compass, and laid a 20 metre transect line
5. A white slate was put in front of the camera, and white balance was set: (Scene>>Auto mode>>Menu>>White Balance>>Measure)
6. The quadrat was dropped at all x values
7. Before the data was logged by photos, a hand sign was used to signify which quadrat it was, on which transect line
8. Three photos were taken of each quadrat at each x value. These photos are how data was logged. The date and GPS location in the EXIF data was used to separate sites out, and the handsign to keep each quadrat separate.
9. Each survey site consists of 5 transects randomly oriented transect lines, each with 6 quadrats (data points). There are five survey sites, very roughly 300 metres apart.[1]

[1]600 metres from origin is the rough mooring area for tourist snorkelling operations, which add nutrients to the area in the form of dog food.

Methods of Analysis

1. Photos were organised first by site, then by transect, then by quadrat in folders on a computer
2. The most visible photo was identified and marked
3. A transparent binder sheet had a grid drawn on (10x10), and was pressed against the computer screen. Positioning was consistent.
4. The photo was opened and aligned with the grid, and the number of 1/100 squares filled by algae was counted and marked
5. These were recorded in an excel table, with average values calculated per transect line, then per site

Variables

In an effort to ensure that only the independent variable (distance from the oil spill) was changing, multiple variables were controlled. The two key variables were depth and geographic location of the survey.

Controlled Variable	**Justification and control methods**
[illegible]	Current has the potential to prevent nutrient sources from collecting, forming the conditions necessary for an algal bloom. Therefore, current has been managed by controlling for depth and region of the island, both of which could change current conditions. If uncontrolled, it could lead to differential nutrient concentrations which could potentially affect rate of algal growth.
[illegible]nt Sources	Given that the investigation aims to quantify the effect of the oil on the algae, any other factors that could potentially affect algae populations must be minimised. Other nutrient sources, like fertiliser runoff, would skew the data in one sample site. Nutrient richness is being controlled by keeping depth constant (there may be different nutrient sources at different levels) and keeping sampling region constant (there may be different nutrient sources in different geographic areas, like ocean upwelling where nutrients are pushed up from the deep ocean by currents). If uncontrolled, additional nutrient sources could lead to excess rates of algal growth and not provide accurate representations of general algal cover in the region.
[illegible] tensity	Light intensity, a key determinant in algae propagation, must be controlled by the same two factors as the previous two variables: depth and location. Different locations could have different light intensity, perhaps due to different amount of sediment in the water or more intense shading from trees above water. Different depths would definitely have not only different light intensities, but different wavelengths of light altogether, as water absorbs light, by increasing wavelength (red, the lowest wavelength, is filtered out first)If uncontrolled, it could lead to algal growth conditions which could potentially heighten algal growth rates without excess nutrients.
Boundary decisions (U [illegible] 2009)	In quadrat surveying, a significant source of error comes from where the quadrat is dropped and how partial objects are counted. An object might be ¾ outside the frame, but clearly that ¼ inside does exist. If rules are clear (eg ½ object inside frame will be counted as 1) then error is minimised. For this experiment, ½ of one transparency box used will be equal to a full percentage cover [illegible] to control.

Dependent Variable

Percentage coverage, measured using a transparent grid overlaid [illegible] were to be rejected and a positive correlation were to be established between proximity to Clifton Pier and percentage algae coverage, percent coverage would change with manipulation of the independent variable (distance). Changing distance would theoretically change percentage cover, by virtue of changing concentration of oil, which according to Birkeland, Reimer and Young (1976) is positively correlated with algal growth, though this has not been established. As the method for quantifying percentage coverage involves counting, uncertainty is expressed through standard deviation.

Independent Variable

Distance from origin of fuel leak, measured in metres from leak. The null hypothesis would be rejected if manipulation of the distance from fuel origin changed percentage cover of algae. Method of manipulation of independent variable is increasing distance from fuel origin by divers swimming and laying transect lines further west. Uncertainties associated with use of metres is ±.001 metres (½ the smallest unit on the transect line).

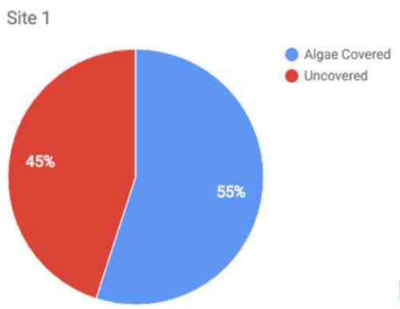
Data

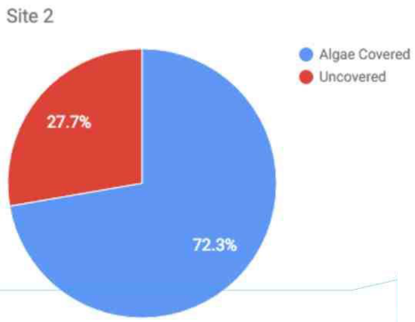
Distance from point source pollution vs percentage algal coverage

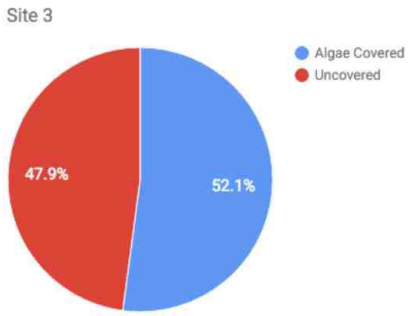
Distance from point source pollution		**A (% cover)**	**B (% cover)**	**C (% cover)**	**D (% cover)**	**E (% cover)**	**F (% cover)**	**transect avg**	**site avg**
1: 300 metres									**55.(**
	transect 1	46.0	54.0	71.0	37.0	27.0	46.0	46.8	
	transect 2	58.0	68.0	42.0	95.0	53.0	45.0	60.2	
	transect 3	37.0	50.0	60.0	49.0	60.0	47.0	50.5	
	transect 4	43.0	35.0	35.0	45.0	69.0	99.0	54.3	
	transect 5	56.0	51.0	94.0	65.0	64.0	49.0	63.2	
2: 600 metres									**72.3**
	transect 1	61.0	76.0	49.0	69.0	52.0	unsure	61.4	
	transect 2	83.0	83.0	55.0	86.0	58.0	32.0	66.0	
	transect 3	68.0	38.0	99.0	99.0	76.0	46.0	71.0	
	transect 4	80.0	38.0	95.0	89.0	80.0	91.0	78.8	
	transect 5	86.0	83.0	75.0	78.0	93.0	90.0	84.2	
3: 900 metres[2]									**52.1**
	transect 1	50.0	80.0	20.0	35.0	32.0	24.0	40.2	
	transect 2	65.0	74.0	68.0	70.0	80.0	16.0	62.2	
	transect 3	25.0	8.0	15.0	8.0	80.0	80.0	36.0	
	transect 4	50.0	90.0	60.0	50.0	80.0	42.0	62.0	
	transect 5	42.0	74.0	90.0	70.0	25.0	60.0	60.2	
4: 1200 metres									**77.0**
	transect 1	75.0	80.0	75.0	83.0	53.0	70.0	72.7	
	transect 2	81.0	88.0	96.0	93.0	87.0	92.0	89.5	
	transect 3	42.0	81.0	59.0	86.0	63.0	64.0	65.8	
5: 1500 metres									**81.2**
	transect 1	88.0	75.0	83.0	70.0	87.0	72.0	79.2	
	transect 2	79.0	75.0	85.0	82.0	96.0	66.0	80.5	
	transect 3	94.0	95.0	68.0	84.0	67.0	95.0	83.8	

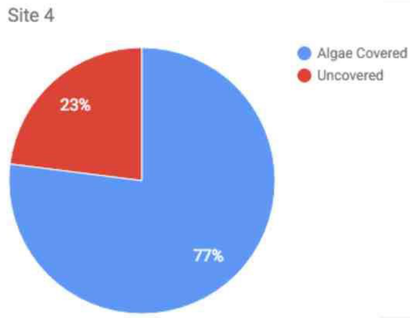
Site with distance from point source	**Average % Coverage**	**Standard Deviation**
1:300 metres	55.0	±17.6
2: 600 metres	72.3	± 19.4
3: 900 metres	52.1	± 26.1
4: 1200 metres	77.0	±14.9
5: 1500 metres	81.2	±10.2

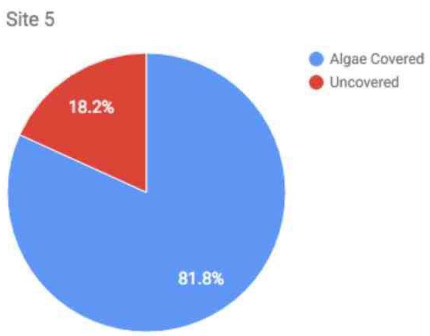
[2] This is also approximately the site where dog food (a source of nutrients) is regularly thrown into the water to attract fish to snorkellers.


Site 1
Algae Covered
Uncovered
45%
55%


Site 2
Algae Covered
Uncovered
27.7%
72.3%

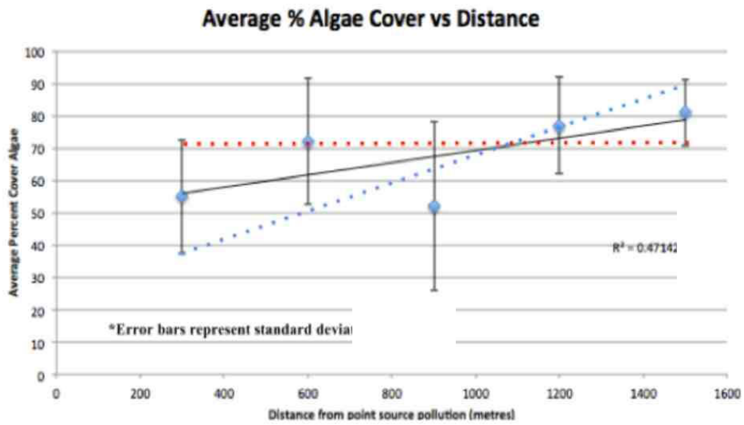

Site 3
Algae Covered
Uncovered
47.9%
52.1%


Site 4
Algae Covered
Uncovered
23%
77%


Site 5
Algae Covered
Uncovered
18.2%
81.8%

Analysis

Any line that can be drawn through all error bars is a possible slope; the black line is a computer simulated line of best fit. As the red line is a potential slope with a correlation coefficient of 0, there is statistical reason to accept the null hypothesis.



The minimum percent coverage of all transects is 36%, though the minimum average of a single survey site is 52% . ' ;e of averages among all 122 quadrats is 53.5, which is large, due to the high variability of the quadra lard deviation is large, too, as a result of the large variation. st percentage cover occurs at 1500 metres, with the lowest at 300 metres. Though the trendline is positive, suggesting algal cover increases with distance, percentage cover at 600 metres, site of nutrient enrichment, is higher than that at 900 metres, further suggesting that injection of nutrients through feeding fish is positively correlated to algal coverage.

Anecdotally, there appears a weak inverse correlation between algal cover and proximity to the site, suggesting more algal growth further away from the point source of pollution. This is suggestive of a rejected null hypothesis; however, statistical analysis does not support this assertion.

n correlation coefficient test was conducted to find the correlation between the two variables of distance and algal coverage;

$$R= \frac{\Sigma(x-Mx)(y-My)}{\sqrt{\Sigma(x-Mx)^2}\sqrt{\Sigma(y-My)^2}}$$

r = 17130 / √((900000)(694.388))

r = 0.6852

R= correlation coefficient.
x: X Value (distance)
y: Y Values (algal coverage)
M_x: Mean of X Values
M_y: Mean of Y Values
$X - M_x$ & $Y - M_y$: Deviation scores
$(X - M_x)^2$ & $(Y - M_y)^2$: Deviation Squared
$(X - M_x)(Y - M_y)$: Product of Deviation Score

Using socscistatistics.com, r was calculated at a value of 0.6852, suggesting distribution of algae is not independent of distance from the plant, as a score of 1 is a perfect positive correlation. The p value, however, was 0.201712. At a significance level of 0.05, the result is not significant, as in 20% of cases the null hypothesis could be accepted and the alternate rejected. Given the number of variables possibly affecting algal coverage and methodological errors associated with data collection, the p value suggests that the null hypothesis should be accepted pending further research.

Evaluation of Methods

There were multiple errors in experimental methodology that could have potentially led to inaccuracies in the data. Future repeats of the experiment should attempt to avoid them.

Sample sites were not spaced consistently. Though approximately 300 metres apart, this is effectively an educated guess with google maps. Sites should have been spaced exactly 300 metres apart (±.001 metres), and gps coordinates should have been plotted for divers to descend at. Failure to do this may have led to overlap between sites and lack of discreteness between sample sites. Additionally, it makes the experiment harder to replicate.

The initial method of measurement included analysing the photos through calculating the integral of the area covered on the photo in relation to a set reference, using loggerpro. This proved ineffective, and the experiment was analysed using a transparency overlaid on the photos. Algae was difficult to see, and this could have resulted in less algae being identified than was actually present, skewing the data. Boundary decision errors, one of the factors that is difficult to control were exacerbated by poor photo quality and different rugosity (roughness and elevation) of the reef. If the experiment were to be conducted again, algal coverage should have been counted underwater, using a quadrat subdivided into tenths. This would be more accurate and not significantly more time consuming, as three usable photos would not have to be taken of each site. Additionally, the diver can better apply the rules needed to minimise boundary decision errors in three dimensions rather than on a photo, where dimensions are generally unclear (depending on camera angles).

In total, 128 quadrats were analysed to produce five average algal cover values, though each average was not created with an equal number of data points. The first three site averages had thirty quadrats each, which made them laborious to process though accurate. The last two only had fifteen, which means they were potentially less accurate. Only five sites were surveyed, which is the bare minimum to find a correlation. In order to definitively accept or reject the null hypothesis, more intervals should have been surveyed. The high sample size is necessary, however, due to the intense variability of the algal coverage, and a repeat of the experiment should keep a high number of repeats. It would be difficult and would require more diving, especially if coverage was counted underwater, but would result in a clearer trend with a greater sample size.

At a site of confirmed nutrient enrichment (600 metres, through tourist snorkelling) percentage cover is higher. Nutrients are concentrated at 600 metres by the feeding operations, leading to eutrophication. The fact that algal coverage is elevated far from the power plant, coupled with the conclusion of the Birkeland, Reimer and Young (1976) study suggests that nutrients may be being concentrated at the 1200 and 1500 metre sites. Possible explanations for this include current (moving nutrients to an area where they settle) or some other unrecognised geographic feature. Either way, this is a failure to control geographic factors that could affect nutrient concentrations and so is a limitation of the methodology.

Conclusion

The investigation aimed to quantify the relationship between proximity to the power plant (a rough measure of expected exposure to oil) and percentage algae cover. Regardless of any correlation to distance, both qualitative and quantitative data show markedly higher levels of algae than is present in healthy Caribbean ecosystems. Healthy reefs are less than 10% covered with algae (NEPA 2014). The lowest percent coverage in Clifton was 52% ("poor", by NEPA standards) with all four other sites "critical"(NEPA, 2014).

At 600 metres there is a spike in average percent coverage: this corresponds to nutrient enrichment associated with tourist snorkelling, suggesting eutrophication is taking place from that activity at sample site 2. This is consistent with literature on the subject, which states that a concentrated excess of nutrients in a sunlit area will produce algal blooms. (Graham 2010).

With respect to the research question "To what extent does proximity to the Clifton Pier power station positively correlate with higher percentage algae coverage?", anecdotal evidence, Pearson testing initially suggested a positive correlation (with high standard deviation, despite extensive sampling) between distance and algal cover (which is rejection of the null hypothesis). On the other hand, a possible correlation coefficient of the slope is 0 and a weak p value suggest otherwise, that the null hypothesis must be accepted. Manipulation of the independent variable was therefore insufficient to impact the dependent variable in a statistically significant way. This may be due to numerous sampling and methodological errors (like inconsistent intervals between sample sites).

The null hypothesis- that proximity to Clifton pier power plant is unrelated to percentage algal cover- must be accepted pending further research in favour of either alternate, that proximity to Clifton Pier power plant is somehow correlated to percentage coverage.

Acceptance of the null hypothesis was not predicted at the beginning of the experiment, though more data is needed to support any claim of any correlation or lack of correlation between proximity to Clifton Pier and algal coverage. An inverse correlation would run counter to what would was suggested by the Birkeland et Al. (1976) study, that a positive correlation exists between algal growth and bunker c;

In the future, controlled testing of the type done in Birkeland, Reimer & Young (1976) should be carried out, using oil samples collected at Clifton Pier. More detailed transect studies in the same area with larger sample sizes could potentially conclude a study with a more decisive result. Despite the unexpected inverse correlation, rates of algae coverage are anomalously high in Clifton bay, to detrimental effect on the marine environment.

Bibliograph

Birkeland, C., Reimer, A., & Young, J. (1976). *Survey of Marine Communities in in Panama and Experiments Oil* (USA, EPA, Dallas). Balboa: Smithsonian Tropical Research Institute. Retrieved June 15, 2015, f http://nepis.epa.gov/Exe/ZyPDF.cgi/2000WO9S.PDF?Dockey=2000WO9S.PDF

DUJS (2012). *Eutrophication in the Gulf of Mexico: How Midwestern farming practices are creating a 'Dead Zone.* Retrieved from http://dujs.dartmouth.edu/winter-2012/eutrophication-in-the-gulf-of-mexico-how-midwestern- farming-practices-are-creating-a-%E2%80%98dead-zone%E2%80%99#.VosbG5MrKHo

Graham, J. (2010) *Algae: The Good, the Bad and the Ugly.* Retrieved from http://ks.water.usgs.gov/static_pages/studies/water_quality/cyanobacteria/jlgraham-08-10-10.pdf

Hartnell, N. (2012) *Bahamas Targeting 5-6% Tourist Growth In 2013.* Retrieved from http://www.tribune242.com/news/2013/feb/05/bahamas-targeting-5-6-tourist-growth-in-2013/

Jackson, J., Donovan, Mary., Cramer, K., & Lam V. (2014). *STATUS AND TRENDS OF CARIBBEAN CORAL REEFS: 1970-2012.* Retrieved from http://cmsdata.iucn.org/downloads/caribbean_coral_reefs___status_report_1970_2012.pdf

NEPA. (2014) *Coral Reefs of Jamaica, An Evaluation of Ecosystem Healt*h. Retrieved from http://www.nepa.gov.jm/new/media_centre/publications/2013_Coral_Reef_Report_card.pdf

NOAA (2016). *Biodiversity*. Retrieved from http://coralreef.noaa.gov/aboutcorals/values/biodiversity/

NOAA. (2016). *No. 6 Fuel oil (Bunker C) Spills.* Retrieved from http://response.restoration.noaa.gov/oil-and-chemical-spills/oil-spills/resources/no-6-fuel-oil-spills.html

Stangroom, Jeremy. (2016). *Pearson Correlation Coefficient Calculator*. Retrieved from http://www.socscistatistics.com/tests/pearson/Default2.aspx

USGS. (2014) *Eutrophication definitions.* Retrieved from http://toxics.usgs.gov/definitions/eutrophication.html

Appendices

Raw Data

600 metres	a	b	c	d	e	f	avg
transect 1	61	76	49	69	52	unsure	61.400
transect 2	83	83	55	86	58	32	66.000
transect 3	68	38	99	99	76	46	71.000
transect 4	80	38	95	89	80	91	78.830
transect 5	86	83	75	78	93	90	84.160
avg							72.270

300 metres	a	b	c	d	e	f	avg
transect 1	46	54	71	37	27	46	46.833
transect 2	58	68	42	95	53	45	60.167
transect 3	37	50	60	49	60	47	50.500
transect 4	43	35	35	45	69	99	54.333
transect 5	56	51	94	65	64	49	63.167
							55.000

900 metres	a	b	c	d	e	f	avg
transect 1	50	80	20	35	32	24	40.160
transect 2	65	74	68	70	80	16	62.160
transect 3	25	8	15	8	80	80	36.000
transect 4	50	90	60	50	80	42	62.000
transect 5	42	74	90	70	25	60	60.167
							52.097

1200 metres		a	b	c	d	e	f	avg	standard deviation
	transect 1	75	80	75	83	53	70	72.667	
	transect 2	81	88	96	93	87	92	89.500	
	transect 3	42	81	59	86	63	64	65.833	
								77.000	±14.852

1500 metres									standard deviation
	transect 1	88	75	83	70	87	72	79.167	
	transect 2	79	75	85	82	96	66	80.500	
	transect 3	94	95	68	84	67	95	83.833	
							site avg	81.177	±10.199