Waikato Regional Council Technical Report 2018/29

Phycocyanin sensor calibration in the Waikato

www.waikatoregion.govt.nz
ISSN 2230-4355 (Print)
ISSN 2230-4363 (Online)

Waikato
REGIONAL COUNCIL
Te Kaunihera ā Rohe o Waikato

Prepared by:
M Stewart and N Phillips
(Streamlined Environmental Ltd)

For:
Waikato Regional Council
Private Bag 3038
Waikato Mail Centre
HAMILTON 3240

October 2018

Document #: 13099057

Peer reviewed by:
Mark Hamer
Deniz Özkundakci

Date October 2018

Approved for release by:
Ed Brown

Date November 2018

Disclaimer

This technical report has been prepared for the use of Waikato Regional Council as a reference document and as such does not constitute Council's policy.

Council requests that if excerpts or inferences are drawn from this document for further use by individuals or organisations, due care should be taken to ensure that the appropriate context has been preserved, and is accurately reflected and referenced in any subsequent spoken or written communication.

While Waikato Regional Council has exercised all reasonable skill and care in controlling the contents of this report, Council accepts no liability in contract, tort or otherwise, for any loss, damage, injury or expense (whether direct, indirect or consequential) arising out of the provision of this information or its use by you or any other party.

Phycocyanin sensor calibration in the Waikato

Action	Name	Date
Draft prepared by	Mike Stewart and Ngaire Phillips	18th September 2018
Draft reviewed by	Mark Hamer and Deniz Özkundakci (WRC)	28th September 2018
Final prepared by	Mike Stewart	23rd October 2018

Lake Whangape Buoy (photo by Mike Stewart)

Stewart, M. and Phillips, N. (2018) Phycocyanin sensor calibration in the Waikato. Report WRC1702–Final, Streamlined Environmental, Hamilton, 76 pp.

Streamlined Environmental Ltd
Hamilton, New Zealand
www.streamlined.co.nz info@streamlined.co.nz

Contents

Executive Summary 1

Methods 1

Regression analysis summary 2

Discussion summary 3

Recommendations 3

1. Introduction 5

2. Methods 7

2.1 Water sampling 7

2.2 Laboratory analyses 9

2.3 Sensor data 9

2.4 Data manipulation 10

2.5 Statistical analyses 10

3. Biotic and abiotic water measurement results 12

3.1 Algae cell counts 12

3.2 Long term blue-green algae cell counts 16

3.3 Blue-green algae genus and species distribution 19

3.4 "Non" blue-green algae genus distribution 22

3.5 Sensor data 23

3.6 Water measurements 27

4. Regression analyses 32

4.1 Site-specific regression analysis of sensor PC fluorescence and blue-green algae cell counts 32

4.2 Site-specific regression analysis of sensor Chl-*a* fluorescence and "non" blue-green algae cell counts 33

4.3 Site-specific regression analysis of sensor Chl-*a* fluorescence and water Chl-*a* concentration 35

4.4 Site-specific regression analysis of laboratory pH measurement and pH from on-site hand-held meter 37

4.5 Site-specific regression of sensor turbidity and laboratory suspended solids 39

5. Discussion 43

5.1 Literature review on PC sensor fluorescence as a tool for cyanobacteria monitoring 43

5.2 Factors influencing phycocyanin sensor effectiveness 45

6. Recommendations 48

7. Acknowledgements 49

8. References 50

Appendix 1 Pivot Graphs 51

Appendix 2 Regression graphs for blue-green algae cell counts vs PC sensor fluorescence . 62

Appendix 3 Regression graphs for "non" blue-green algae cell counts vs Chl-*a* sensor fluorescence 67

Appendix 4 Regression graphs for turbidity (sensor) vs turbidity (WQ meter) 72

Executive Summary

Blooms of cyanobacteria (or blue-green algae) have the potential to introduce into the water toxins that can have acute and, potentially, fatal consequences for the public.

Most cyanobacterial monitoring includes analysis by conventional laboratory methods such as taxonomic analysis (cell counts and biomass measurements), phytoplanktonic pigment extractions and cyanotoxin analysis of water samples. These methods are costly, time consuming and are unable to detect rapid changes in water quality or sudden increases of cyanobacterial biovolume.

For most freshwater cyanobacteria, phycocyanin (PC) is a typical pigment, which is unique for its fluorescence properties, which has a different emission range to chlorophyll a (Chl-a), the major pigment from eukaryotic algae (green, red and brown algae).

Online fluorescence-based probes can exploit this difference *in vivo* fluorescence emission ranges of PC and Chl-*a* and may be able to estimate populations of blue-green algae. Furthermore, real-time fluorescence readings may provide immediate warnings of blue-green algae blooms. However, robust and consistent relationships between PC fluorescence readings and blue-green algae population measures are necessary before sensor readings could replace conventional laboratory measurement of algae cell count and biovolume.

Waikato Regional Council (WRC) has continuous monitoring stations in 3 lakes[1] (Waikare, Whangape, and Waahi) and 2 Waikato River sites (Elbow and Hamilton Traffic). Through river sonde and lake buoy sensors, these stations measure water Chl-*a* and PC fluorescence, plus dissolved oxygen (DO), water temperature, pH and turbidity. WRC are interested in understanding the relationships between the buoy sensor readings and either water quality meter readings or laboratory measurements from lake water samples, collected near the buoys. Principally, this involved examining site-specific relationships between:

- PC sensor (in relative fluorescence units – RFU) and laboratory blue-green algae cell counts (cells/mL);
- Chl-*a* sensor fluorescence (RFU) and laboratory non-blue-green algae cell counts (cells/mL);
- Chl-*a* sensor fluorescence (RFU) and laboratory Chl-*a* water concentration (mg/m^3);
- pH from water quality meter and pH measured in the laboratory;
- Turbidity sensor readings (nephelometric turbidity unit – NTU) and laboratory suspended solids measurements (g/m^3).

Methods

Lake sampling was carried out monthly from April 2017 to May 2018 with sampling alternated between WRC and Streamlined Environmental Ltd (SEL). River sampling was undertaken between June 2017 and June 2018, with all sampling undertaken by WRC. All water samples were collected by boat with the exception of the Waikato river at elbow site which was collected from the jetty.

[1] A fourth lake – Ngaroto – was added after this study commenced.

To be consistent with sensor readings, water samples were collected once at the time that the sensor was collecting data.

Lugol's fixed water samples were analysed for algae to species level by NIWA Algal Services. All algae results were reported as cell counts (cells/mL). Water samples were analysed for Chl-*a*, pH and suspended solids (SS) by NIWA Hamilton Water Quality Laboratory.

Median lake sensor data (n=6) were obtained from WRC HydroTel database. Single river sensor data (or an average of 2 datapoints if water collection was outside the sensor output time) were supplied by WRC. SEL field data were entered manually, while WRC field data were supplied by WRC.

Data were supplied by NIWA in Excel format. For each site, algae were separated into blue-green and "non" blue-green algae species. Pivot Tables were used to summarise each.

Regression analyses were undertaken in Excel, with relationships reported as the coefficient of determination (R^2) and significance (P-value). A P-value (*P*) of <0.05 was considered significant.

For the purposes of describing relationships, we have created a relative narrative scale to describe the coefficient of determination: R^2 = 0.80 – 1.00 (strong relationship); R^2 = 0.60 – 0.79 (moderately strong relationship); R^2 = 0.40 – 0.59 (moderate relationship); R^2 = 0.20 – 0.39 (moderately weak relationship); and R^2 = 0.00 – 0.19 (weak relationship).

Regression analysis summary

Coefficient of determination (R^2) for biotic indices (blue-green and "non" blue-green algae cell counts) with applicable sensor fluorescence data (phycocyanin and Chl-*a*, respectively) were generally weak, unless the data were pooled to increase the statistical power of the regression analysis. Conversely, Chl-*a* sensor fluorescence and abiotic indices (hand-held water quality measurement of pH and turbidity) generally had markedly stronger relationships with appropriate laboratory measurements (Chl-*a*, pH, SS, respectively), despite lower number of datapoints.

Pooling of data generally improved the strength and significance of relationships, suggesting greater statistical robustness requires more datapoints than generally obtained for individual sites (n<15), at least for biotic indices. However, this is not always the case and confounding factors are present which lead to variation of data. Data outliers had a marked effect on strength and significance of some relationships, especially for Chl-*a* sensor fluorescence relationship with "non" blue-green algae cell counts.

Non-linear regressions (logarithmic, polynomial, power) were also briefly investigated but these showed no consistent improvement to linear regressions. Furthermore, they are inconsistent with linear relationships reported in the literature.

For algae regressions, drilling down to genus and/or species taxonomic level generally reduced the strength of R^2. This was partly due to different dominant genus/species between sites and sporadic occurrence of many genus/species over the 14-month timeframe.

Discussion summary

The linear regressions between PC fluorescence and blue-green algae cell counts noted in the current study are consistent with literature. Comparisons of the strength of these relationships between the current study and literature are confounded due to complexities involved. These include:

- Different methodologies;
    - blue-green algae enumeration by cell count and/or biovolume, or PC fluorescence from extracted blue-green algae;
    - different sensor precision and sensitivity;
    - number of datapoints;
    - use (or not) of thresholds to deal with "upper" and "lower" bound sensor fluorescence issues;
    - experimental bias.
- Varying external effects from biotic and abiotic factors;
    - Different site and season specific algae profile;
        - Species-specific PC fluorescence;
        - False-positives from "non" blue-green algae (especially diatoms and green algae) when blue-green algae are present in low proportion of total algae;
    - Genetic and life stage variability of PC production in blue-green algae;
    - Algae heterogeneity (aggregation);
    - Varying and lesser known abiotic effects, specifically turbidity, temperature, and light.

Recommendations

This study has set the foundation for complementing water sampling and enumeration of blue-green algae cell counts and/or biovolume with phycocyanin sensor fluorescence.

However, there needs to be more information gathered around biotic and abiotic interferences, and their impact at each site. Also, a greater understanding of blue-green algae and "non" blue-green algae species dynamics is necessary for each site. This would need to include consistency of any seasonal variation and potential spatial variation at lake sites and different river reaches. More information is needed on sensor capabilities and accuracy.

Setting upper and lower sensor thresholds may be necessary to improve the variability. However, there are site and sensor specific logistical issues with this.

Recommendations for the future include:

- Extending the dataset at the current sites to increase the number of datapoints and statistical power. This will also allow for assessment of site-specific stability of algae species throughout repeated seasons;
- At critical times of the year – specifically in times of high and low blue-green algae cell counts – increase the number of water monitoring events;
    - For high cell counts – and restricted primarily to Waikare and Whangape where blooms are frequent and of high intensity – this would be useful for assessing a

different sensor brand and/or configuration during periods where sensor saturation may be occurring;
    - For low cell counts where negative values are consistently encountered, i.e. specifically for Waahi, assess whether an EXO sonde arrangement (as for river sites which have low blue-green algae cell counts but no negative fluorescence results) improves the reliability of data;
- Adding new sites of interest, especially lake sites with known issues of blue-green algae blooms;
- Including biovolume data. There is significant variability of blue-green algae sizes and biovolume will normalise these data. Furthermore, with NPS-FM guidelines based on biovolume, this will provide a more direct link between sensor fluorescence and regulatory thresholds.

1. Introduction

Over recent years, water bodies in some parts of New Zealand have experienced an increase in the number of cyanobacterial blooms. These events have the potential to introduce into the water toxins that can have acute and, if their concentrations are high enough, fatal consequences for consumers (Ministry of Health, 2018).

Cyanobacteria are primarily aquatic organisms with many characteristics of bacteria. As their metabolism is based on photosynthesis, they have also been termed blue-green algae.[2] They may grow as filaments or colonies readily visible and identified (to the genus level) under a microscope. Cyanobacteria are not, of themselves, a health hazard, but the toxins they produce (called cyanotoxins) are (Ministry of Health, 2018).

Most cyanobacterial monitoring includes analysis by conventional laboratory methods such as taxonomic analysis (cell counts and biomass measurements), phytoplanktonic pigment extractions and cyanotoxin analysis of water samples. These methods are costly, time consuming and are unable to detect rapid changes in water quality or sudden increases of cyanobacterial biovolume (Zamyadi et al., 2012b).

In New Zealand, the National Policy for Freshwater Management sets thresholds (numeric attribute states) to protect human health from cyanobacteria during recreational activities based on biovolume (Ministry for the Environment, 2017).

For most freshwater cyanobacteria, phycocyanin (PC) is a typical pigment, which is unique for its fluorescence properties. While fluorescence of eukaryotic algae is effectively excited by blue light (approximately 430–530nm) with an emission peak around 685nm – the fluorescence emission maximum of chlorophyll *a* (Chl-*a*) – cyanobacterial PC is excited in the orange and red parts of the spectrum (approximately 590–630nm) and its emission maximum lies around 650 nm (Gregor et al., 2007).

Online fluorescence-based probes can exploit the different *in vivo* fluorescence emission ranges of PC and Chl-*a* and may be able to estimate populations of blue-green algae. Furthermore, real-time fluorescence readings may provide immediate warnings of blue-green algae blooms. However, robust and consistent relationships between PC fluorescence readings and blue-green algae population measures are necessary.

Waikato Regional Council (WRC) has continuous monitoring stations in 3 lakes[1] (Waikare, Whangape, and Waahi) and 2 Waikato River sites (Elbow and Hamilton Traffic) (see Figure 1). Through river sonde and lake buoy sensors, these stations measure water Chl-*a* and PC fluorescence, plus dissolved oxygen (DO), water temperature, pH and turbidity. WRC are interested in understanding the relationships between the buoy sensor readings and either water quality meter readings or laboratory measurements from lake water samples, collected near the buoys. Principally, this involved examining site-specific relationships between:

- PC sensor (in relative fluorescence units – RFU) and laboratory blue-green algae cell counts (cells/mL);

[2] In this report we primarily use the term blue-green algae.

- Chl-*a* sensor fluorescence (RFU) and laboratory non-blue-green algae cell counts (cells/mL);
- Chl-*a* sensor fluorescence (RFU) and laboratory Chl-*a* water concentration (mg/m^3);
- pH from water quality meter and pH measured in the laboratory;
- Turbidity sensor readings (nephelometric turbidity unit – NTU) and laboratory suspended solids measurements (g/m^3).

2. Methods

2.1 Water sampling

Information on site location and sensor configuration is summarised in Figure 1 and Table 1.

Lake sampling was carried out monthly from April 2017 to May 2018 with sampling alternated between WRC and SEL. River sampling was undertaken between June 2017 and June 2018, with all sampling undertaken by WRC (Table 2). In total there were 14 sampling events for Waikare, Whangape[3], Waahi, and Hamilton Traffic and 15 sampling events for Elbow (Table 2).

All water samples were collected by boat. To be consistent with sensor readings, water samples were collected once at the time that the sensor was collecting data (every 15 minutes for lake sites and every 20 minutes for river sites). Water samples were collected at a depth of 0.5 m in lakes and 1.0m in rivers to be consistent with the depth of sensor (Table 1). Water quality parameters – pH, turbidity, temperature, dissolved oxygen (DO) and specific conductivity – were measured at the same time using a YSI Pro-DSS handheld multiparameter water quality meter, supplied by WRC.[4]

Two water samples were collected in high density polyethylene bottles (supplied by NIWA):

1. Samples for algae cell counts were collected in a 250 mL bottle and immediately fixed with Lugol's solution;
2. Samples for all other laboratory analyses were collected in a 2L bottle.

Samples were kept chilled in a chilly bin in the dark and delivered to NIWA Hamilton on the same day that sampling occurred.

[3] Whangape lake water level on 30/1/18 was too low to launch the boat so water samples were collected by wading from shore. The location was not near the buoy. These data were included in biotic and abiotic water measurement summaries as they are still representative of Whangape at that time but were subsequently removed from regression analyses due to significant distance from buoy.

[4] There was one exception: for the May 2017 sampling, SEL used a YSI Pro-ODO meter which did not measure turbidity or specific conductivity.

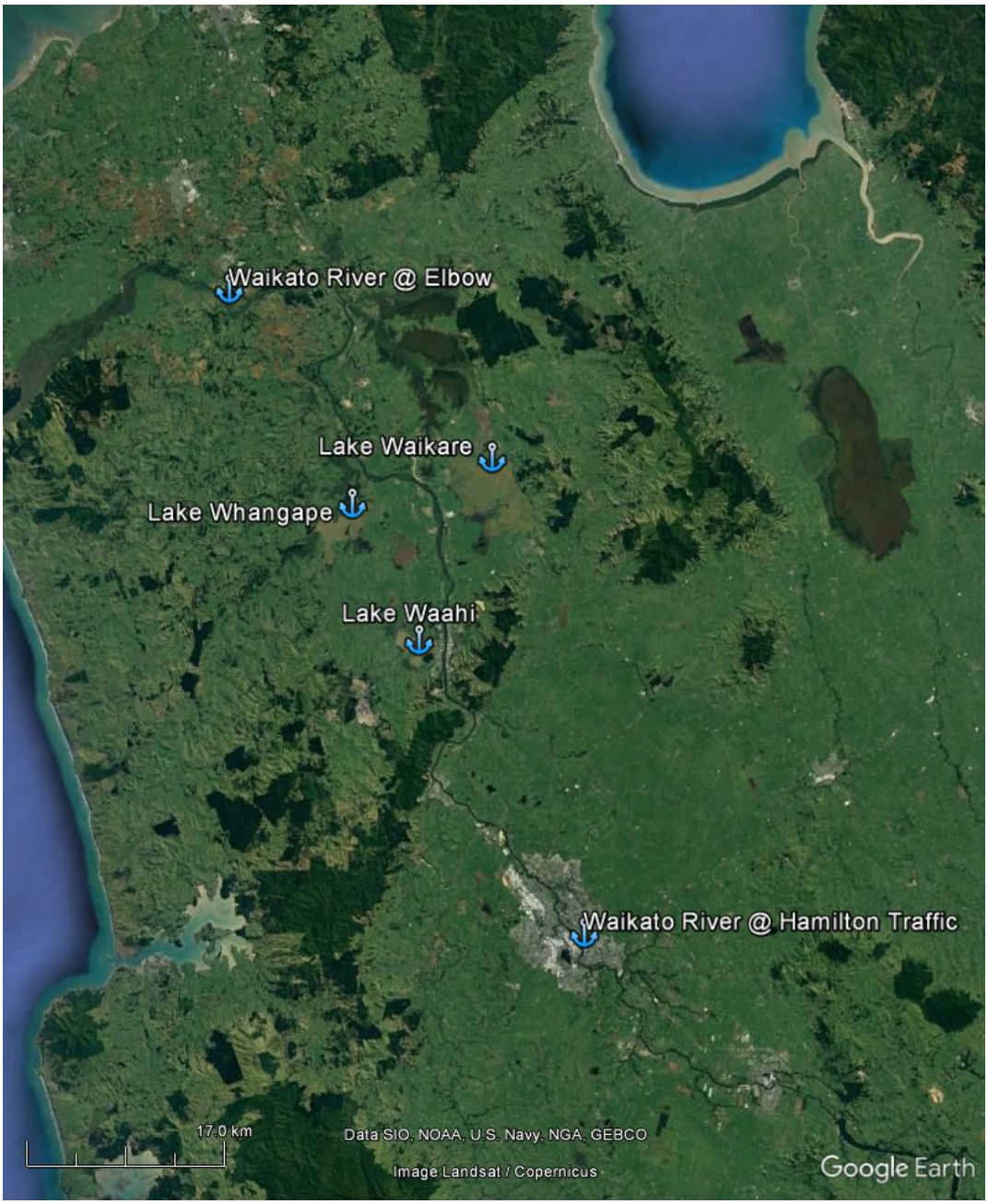


Figure 1. Five river and lake WRC sampling sites included in this study.

Table 1. Site and sensor information.

Site Name	Site Number	Coordinates (NZTM)	Sensor type[1]	Depth
Waikato River @ Hamilton	1131_64	E 1801738 N 5814759	EXO Sonde	1m
Waikato River @ Elbow	1131_133	E 1772410 N 5870516	EXO Sonde	1m
Lake Whangape Buoy	330_14	E 1781993 N 5851747	'Trilux'	0.5m
Lake Waahi Buoy	324_12	E 1788103 N5840152	'Trilux'	0.5m
Lake Waikare Buoy	326_58	E 1794662 N 5855581	'Trilux'	0.5m

[1] Sensors are as per Hodges et al. (2018); Chelsea Technologies LTD 'Trilux' for lake sites and YSI EXO Sonde for river sites.

Table 2. Date sampling undertaken and by whom[1]

Month[2]	Waikare	Whangape	Waahi	River@Elbow	River@Hamilton
Apr-17	24/04/2017	24/04/2017	24/04/2017	No sampling	No sampling
May-17	24/05/2017	24/05/2017	24/05/2017	7/06/2017	1/06/2017
Jun-17	29/06/2017	29/06/2017	29/06/2017	3/07/2017	4/07/2017
Jul-17	25/07/2017	25/07/2017	25/07/2017	10/08/2017	10/08/2017
Aug-17	24/08/2017	24/08/2017	24/08/2017	5/09/2017	5/09/2017
Sep-17	27/09/2017	27/09/2017	27/09/2017	28/09/2017	29/09/2017
Oct-17	26/10/2017	26/10/2017	26/10/2017	31/10/2017	6/11/2017
Nov-17	4/12/2017	4/12/2017	4/12/2017	30/11/2017	18/12/2017
Dec-17	21/12/2017	21/12/2017	21/12/2017	20/12/2017	3/01/2018
Jan-18	30/01/2018	30/01/2018[3]	30/01/2018	3/01/2018	19/02/2018
Feb-18	22/02/2018	22/02/2018	22/02/2018	5/02/2018	26/02/2018[4]
Mar-18	26/03/2018	26/03/2018	26/03/2018	26/02/2018	4/05/2018
Apr-18	27/04/2018	27/04/2018	27/04/2018	4/04/2018	28/05/2018
May-18	25/05/2018	25/05/2018	25/05/2018	2/05/2018	6/06/2018
Jun-18	No sampling	No sampling	No sampling	6/06/2018	19/06/2018
Jul-18	No sampling	No sampling	No sampling	19/06/2018	No sampling
N	14	14	14	15	14

[1] WRC highlighted blue; SEL highlighted green.
[2] Arbitrary month of sampling.
[3] Whangape lake water level on 30/1/18 was too low to launch the boat so water samples were collected by wading from shore. The location was not near the buoy. These data were included in biotic and abiotic water measurement summaries as they are still representative of Whangape at that time but were subsequently removed from regression analyses due to significant distance from buoy.
[4] Water samples were received but no field data were logged on this date.

2.2 Laboratory analyses

Lugol's fixed water samples were analysed for algae to species level by NIWA Algal Services in accordance with NIWA SOP#1-6; Microscopic analysis of settled sample. All algae results were reported as cell counts (cells/mL).

Water samples were analysed for Chl-*a*, pH and suspended solids (SS) by NIWA Hamilton Water Quality Laboratory in accordance with methods A*10200H, APHA 4500H, APHA 2540D, respectively. pH was only measured from December 2017.

2.3 Sensor data

Lake sensors take a reading every 15 minutes at 0, 15, 30, and 45 minutes past each hour. A reading is recorded over 10 seconds and this is repeated 6 times over a minute. Average and median data are then recorded. Data used here are median data over the 1-minute period.

River sondes take a single 10 second reading every 20 minutes at 0, 20, and 40 minutes past each hour. Data used are single values – if collection was at the same time as the sensor output – or an average of two values if water collection fell between two sensor values.

2.4 Data manipulation

Phytoplankton data were supplied by NIWA in Excel format. Data for each site were pooled into a separate tab (one for algae data and one for water quality parameters), with the raw spreadsheet data retained as separate tabs.

For each site, algae were separated into blue-green and "non" blue-green algae species. Pivot Tables were used to summarise each. To avoid addition of data when samples were collected in the same month, samples were summarised by sample # (Table 4). Pivot Graphs are presented in Appendix 1 at genus level.[5]

SEL field data were entered manually into the appropriate water quality tab. WRC field data were supplied in Excel format and copied and pasted into the appropriate water quality tab.

Lake sensor data were obtained from the WRC Hydrotel database through an external connection service and entered manually. Both "average" and "median" sensor data were extracted, however only median data were used for regression analyses.[6] River sensor data were supplied by WRC in Excel format.

2.5 Statistical analyses

Regression analyses were undertaken in Excel, with relationships reported as the coefficient of determination (R^2) and significance (P-value). A P-value (*P*) of <0.05 was considered significant. Although R^2 explains the fit of the regression line to the data, there are no established guidelines for determining what an acceptable value is, as it depends on the data and applications that the data are being used for. For the purposes of describing relationships, we have created a relative narrative scale to describe the coefficient of determination (Table 3).

Table 3. Relative narrative scale to describe the coefficient of determination.

Coefficient of determination (R^2) scale	Relative narrative for relationship
0.80 – 1.00	Strong
0.60 – 0.79	Moderately strong
0.40 – 0.59	Moderate
0.20 – 0.39	Moderately weak
0.00 – 0.19	Weak

For algae comparisons with PC and Chl-*a* sensor data, regressions were undertaken for total algae cell counts initially. For blue-green algae, regressions were also undertaken for dominant[7] and prevalent[8] genus and species. For "non" blue-green algae species regressions were only

[5] Species level added a level of complexity to the graphs that lost clarity of viewing.
[6] The TriLux sensor powers up every 15 minutes and the datalogger records a 'batch' of 5 readings from the sensor. The standard variable (without the _MED suffix) is the **average** of those 5 readings, whereas the variable with the _MED suffix is the **median** of those 5 readings.
[7] Dominant genus and species based on cell counts.
[8] Prevalent genus and species based on number of times detected out of 14/15 sampling events.

undertaken to dominant and prevalent genus.[9] Where a genus or species was not present the data were zero-filled for statistical regression analyses.

Graphs of regression analyses are shown in Appendix 2 to 4.

[9] In practice genus taxonomic level was sufficient as there was generally only 1 species for each genus. The exceptions to this were *Flagellates/Unicells* (predominantly <5µm species) and *Oocystis* (predominantly *Oocystis* sp.).

3. Biotic and abiotic water measurement results

3.1 Algae cell counts

3.1.1 Lake sites

Algae cell counts for lake sites are summarised by sampling event in Table 4 and Figure 2.

Total algae cell counts have been separated into blue-green and “non” blue-green algae cell counts, from which a “monthly” percentage of blue-green algae of total algae was calculated.

Over the 14 sampling events, Waikare was dominated by blue-green algae, with a range of 77-99% of total algae cell counts. Whangape algae profile was dominated by “non” blue-green algae from April to early December 2017 (67-99% “non” blue-green algae), however there was a rapid flip to blue-green dominance in late December 2017 until the end of sampling in late May 2018 (87-99% blue-green algae). Waahi algae profile was dominated by “non” blue-green algae for the majority of the sampling period (71-100% “non” blue-green algae), except for the summer months of late December 2017 to late February 2018 where blue-green algae dominated (88-99% blue-green algae).

The proportion of blue-green algae to total algae is important when considering interferences to the PC sensor. This is discussed in Section 5.1.

Table 4. Lake algae cell counts differentiated by blue-green and "non" blue-green algae.

Date	Sample #	Total blue-green algae	Total "non" blue-green algae	Total algae	Blue-green percentage of total
		Algae cell count (cells/mL)			
Waikare					
24/04/2017	1	1,405,874	56,964	1,462,838	96
24/05/2017	2	431,424	91,710	523,134	82
29/06/2017	3	665,289	34,862	700,151	95
25/07/2017	4	2,799,268	97,830	2,897,098	97
24/08/2017	5	4,303,486	1,273,403	5,576,889	77
27/09/2017	6	471,980	30,299	502,279	94
26/10/2017	7	2,168,634	50,071	2,218,705	98
4/12/2017	8	1,455,282	19,456	1,474,738	99
21/12/2017	9	4,185,220	391,788	4,577,008	91
30/01/2018	10	3,752,787	100,681	3,853,468	97
22/02/2018	11	2,142,939	36,442	2,179,381	98
26/03/2018	12	61,491	9,505	70,996	87
27/04/2018	13	118,394	12,637	131,031	90
25/05/2018	14	595,873	36,985	632,858	94
Whangape					
24/04/2017	1	138	14,483	14,621	1
24/05/2017	2	7,757	16,002	23,759	33
29/06/2017	3	460	6,071	6,531	7
25/07/2017	4	170	1,286	1,456	12
24/08/2017	5	12	1,306	1,318	1
27/09/2017	6	1,110	4,576	5,686	20
26/10/2017	7	220	24,390	24,610	1
4/12/2017	8	48	17,224	17,272	0
21/12/2017	9	1,745,258	269,053	2,014,311	87
30/01/2018	10	263,494	5,374	268,868	98
22/02/2018	11	343,052	9,679	352,731	97
26/03/2018	12	1,201,303	17,475	1,218,778	99
27/04/2018	13	544,076	6,143	550,219	99
25/05/2018	14	92,974	2,684	95,658	97
Waahi					
24/04/2017	1	7	515	522	1
24/05/2017	2	20	393	413	5
29/06/2017	3	28	1,291	1,319	2
25/07/2017	4	59	662	721	8
24/08/2017	5	-	641	641	0
27/09/2017	6	20	1,262	1,282	2
26/10/2017	7	20	549	569	4
4/12/2017	8	921	2,244	3,165	29
21/12/2017	9	70,733	9,513	80,246	88
30/01/2018	10	165,143	1,806	166,949	99
22/02/2018	11	165,438	2,981	168,419	98
26/03/2018	12	78	904	982	8
27/04/2018	13	139	3,310	3,449	4
25/05/2018	14	113	3,052	3,165	4

Figure 2. Lake algae cell counts differentiated by blue-green and "non" blue-green algae. Total cell counts in left column and 100% normalised total cell counts in right column.

3.1.2 River sites

Algae cell counts for river sites are summarised by sampling event in Table 5 and Figure 3. Total algae cell counts have been separated into blue-green and "non" blue-green algae cell counts, from which a "monthly" percentage of blue-green algae of total algae was calculated.

Over the 14 sampling events, Hamilton Traffic was exclusively dominated by "non" blue-green algae (59-100% "non" blue-green algae), except for late February 2018, where algae cell counts

spiked (with a significant blue-green proportion of 41%), cell counts were very low (<7,000 cells/mL). Elbow site had a different algae profile to Hamilton Traffic and lake sites. Blue-green algae dominated in winter (early June to early September 2017) and *not* summer, with 73-98% blue-green algae. After this period, "non" blue-green algae dominated with 76-100% of the total cell counts. The last sampling event (late June 2018), blue-green algae appeared to be increasing (38% of total cell count) and may be following the winter trend from 2017, albeit slightly delayed.

Table 5. River algae cell counts differentiated by blue-green and "non" blue-green algae.

Date	Sample #	Total blue-green algae	Total "non" blue-green algae	Total algae	Blue-green percentage of total
Hamilton Traffic					
1/06/2017	1	15	506	521	3
4/07/2017	2	20	865	885	2
10/08/2017	3	5	463	467	1
5/09/2017	4	15	2,067	2,082	1
29/09/2017	5	6	4,253	4,259	0
6/11/2017	6	24	3,919	3,943	1
18/12/2017	7	223	5,067	5,290	4
3/01/2018	8	174	6,555	6,729	3
19/02/2018	9	82	1,841	1,923	4
26/02/2018	10	4,599	6,586	11,185	41
4/05/2018	11	31	925	956	3
28/05/2018	12	14	1,117	1,131	1
6/06/2018	13	79	1,259	1,338	6
19/06/2018	14	37	939	976	4
Elbow					
7/06/2017	1	6,817	1,347	8,164	84
3/07/2017	2	1,921	695	2,616	73
10/08/2017	3	16,959	344	17,303	98
5/09/2017	4	7,893	1,494	9,387	84
28/09/2017	5	102	4,867	4,969	2
31/10/2017	6	15,383	8,591	23,974	64
30/11/2017	7	53	3,251	3,304	2
20/12/2017	8	882	6,413	7,295	12
3/01/2018	9	-	5,157	5,157	0
5/02/2018	10	934	2,987	3,921	24
26/02/2018	11	57	2,499	2,556	2
4/04/2018	12	140	2,579	2,719	5
2/05/2018	13	26	715	741	4
6/06/2018	14	103	745	848	12
19/06/2018	15	1,175	1,877	3,052	38

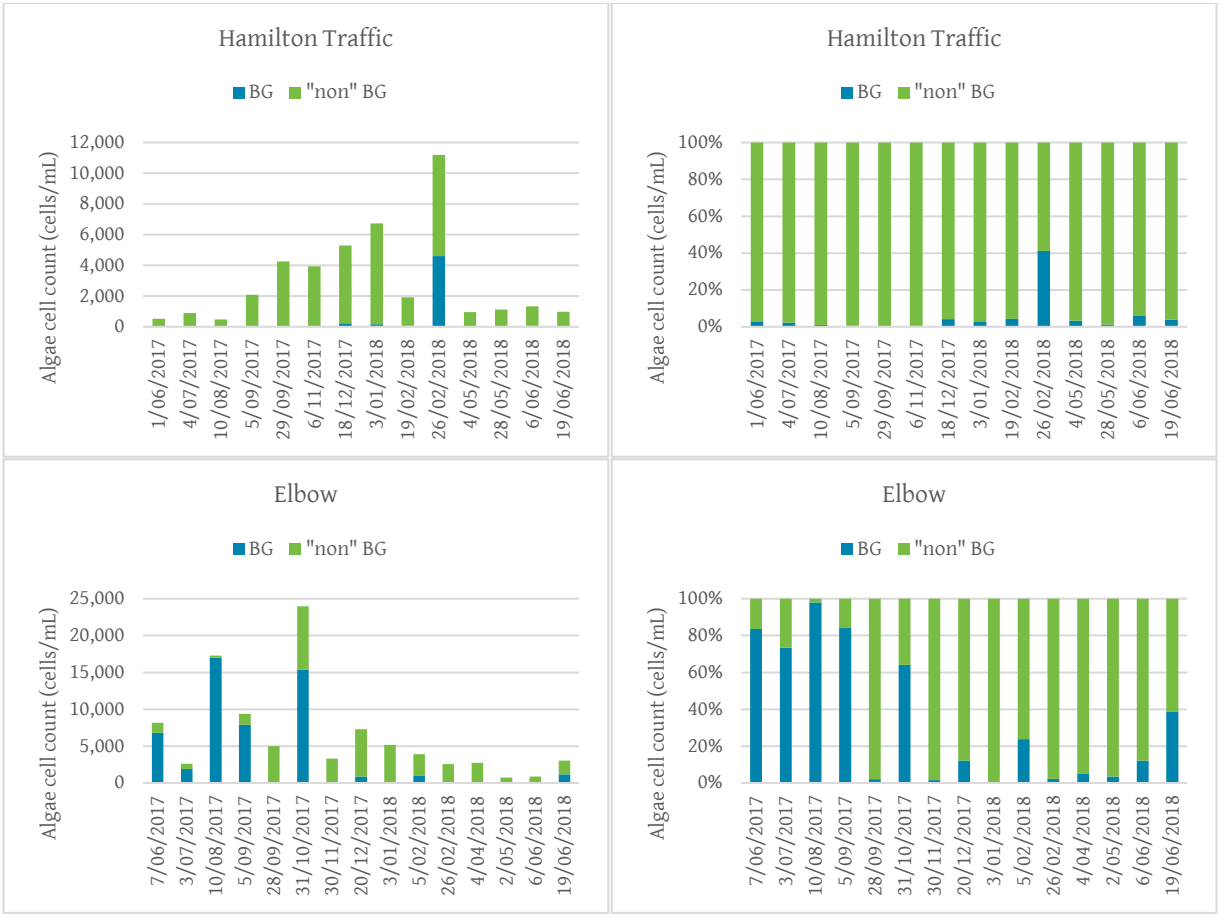


Figure 3. River algae cell counts differentiated by blue-green and "non" blue-green algae. Total cell counts in left column and 100% normalised total cell counts in right column.

To provide perspective, long term blue-green algae lake cell counts are discussed further in Section 3.2.

3.2 Long term blue-green algae cell counts

3.2.1 Lake sites

Blue-green algae cell count data from August 2005 to May 2018 for Waikare, Waahi and Whangape are summarised in Table 6 and Figure 4.

Over the last 13 years, Waikare has consistently experienced blue-green algae cell counts in the 'millions' with a median value of 1.6 million cells/mL and a maximum over 9 million cells/mL. Over the duration of the current project (April 2017 to May 2018) blue-green algae cell counts in Waikare were slightly lower, with a median value of 1.4 million cells/mL and maximum of 4.3 million cells/mL.

Over the last 13 years, Whangape blue-green algae cell counts have been around 10-fold less than Waikare (median ca. 136,000 vs 1,600,000 cells/mL), although in early February 2007 spiked to nearly 11 million cells/mL, considerably higher than the Waikare maximum of 9.1 million

cells/mL. Over the duration of the current project (April 2017 to May 2018), Whangape blue-green algae cell counts have been relatively very low, with a median value of 4,434 cells/mL, which is 3% of the long-term median. Over the latter half of sampling (December 2017 to April 2018) Whangape blue-green algae cell counts increased markedly and ranged from 264,000 to 1.7 million cells/mL.

Over the last 13 years, Waahi blue-green algae cell counts have generally been the lowest of the 3 lakes, with a median value of 50,978 cells/mL. However, between December 2013 and May 2015 Waahi experienced extremely high levels of blue-green algae (Figure 4) with cell counts nearing 15 million cells/mL. Over the duration of the current project (April 2017 to May 2018), Waahi blue-green algae cell counts have been extremely low relatively, with a median value of 69 cells/mL, which is 0.1% of the long-term median.

Table 6. Summary of long-term and short-term blue-green algae cell counts in Waikare, Whangape and Waahi.

Statistic	Waikare	Whangape	Waahi
Long-term (Aug 2005 - May 2018) cell count (cells/mL)			
Median	1,627,058	136,312	50,978
Minimum	37,306	0	0
95th Percentile	6,935,569	3,661,225	7,154,480
Maximum	9,134,668	10,963,931	14,816,316
Short-term (Apr 2017 - May 2018) cell count (cells/mL)			
Median	1,430,578	4,434	69
Minimum	61,491	12	0
95th Percentile	4,226,612	1,391,687	165,246
Maximum	4,303,485	1,745,258	165,438

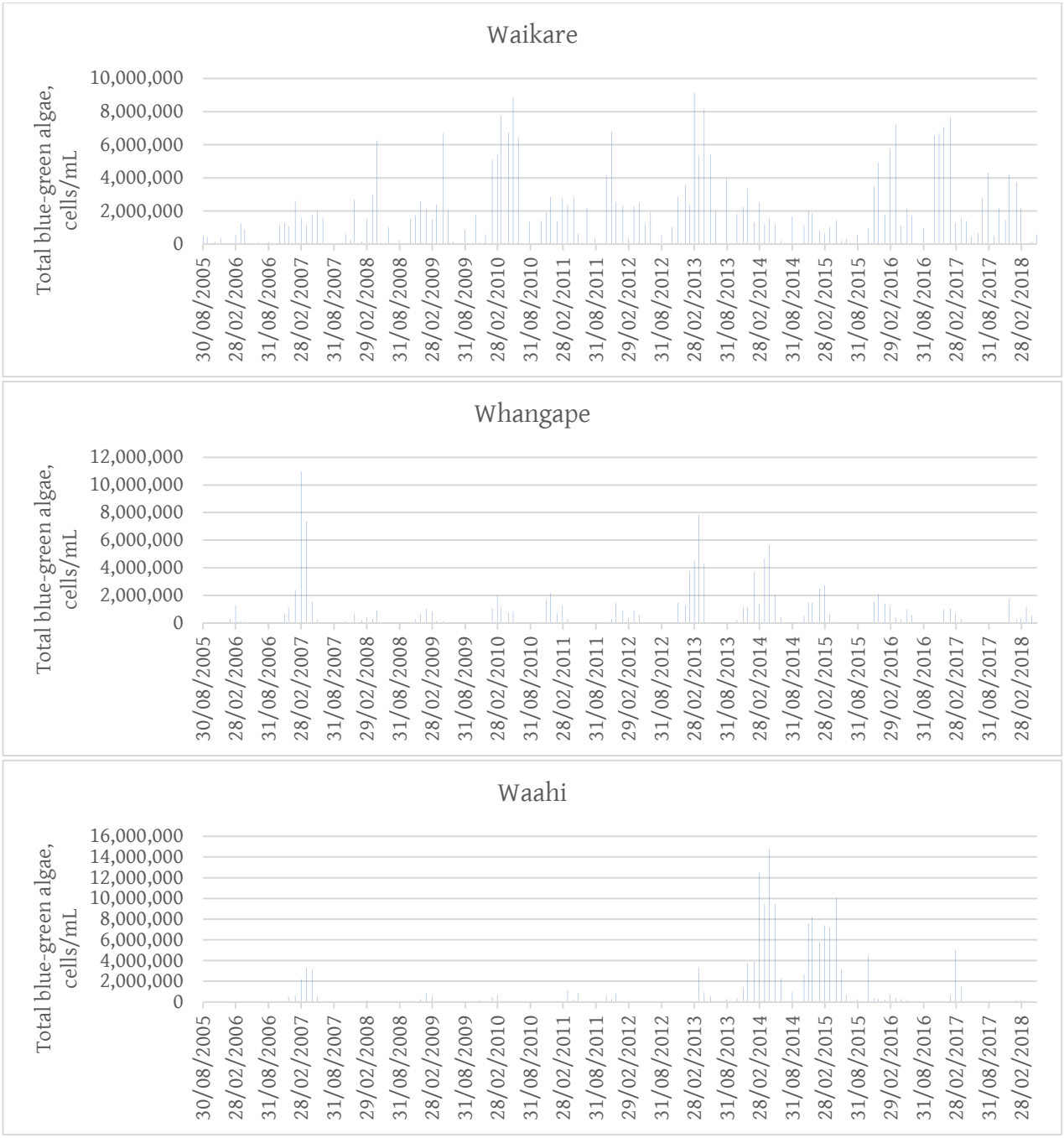


Figure 4. Graphs of long-term blue-green algae cell counts in Waikare, Whangape and Waahi.

3.2.2 River sites

There are no long-term records of blue-green algae data for Hamilton Traffic or Elbow sites. However, Hamilton City Council (HCC) monitor blue-green algae at Wairoa Water Treatment Plant intake, located approximately 500m upstream of Hamilton Traffic site. HCC provided blue-green algae cell counts from June 2011 to August 2018 which are presented in Figure 5. Blue-green algae cell counts at Wairoa intake are markedly lower than the lake sites, with minimum, median, 95th percentile, and maximum cell counts of 0, 67, 6,298 and 17,006 cells/mL, respectively.

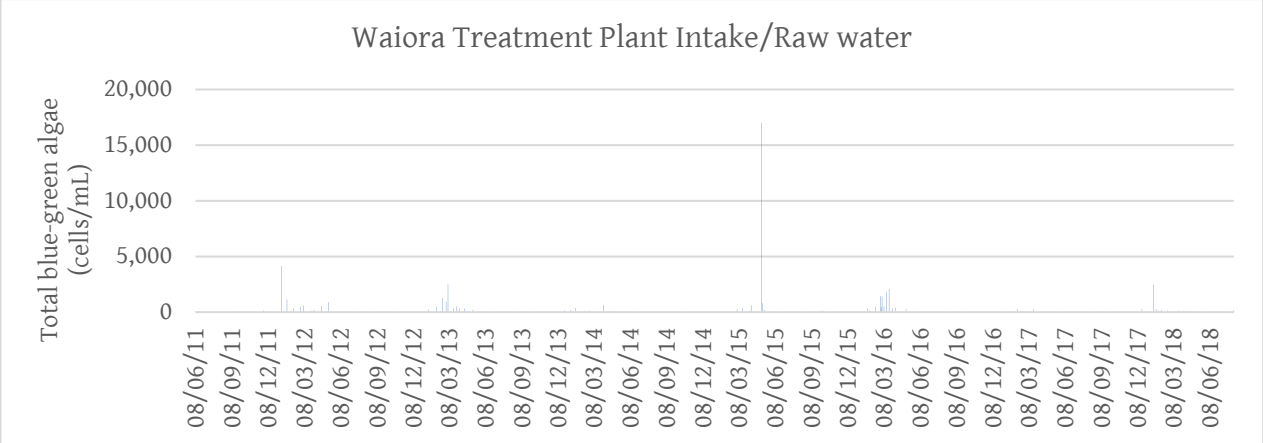


Figure 5. Graph of long-term blue-green algae cell counts for Waikato River site at Wairoa.

3.3 Blue-green algae genus and species distribution

A breakdown of dominant blue-green algae into genus and species level over the sampling period is summarised in Table 7.

Lake Waikare was dominated by 5 blue-green algae genera – *Planktolyngbya*, *Pseudanabaena, Coelomoron, Aphanocapsa*, and *Merismopedia* – accounting for 95% of blue-green algae cell counts over the study duration. Although *Planktolyngbya subtilis* (36%) was the dominant species in terms of cell count, it was only present for 6 of 14 months. *Planktolyngbya* cf. *tallingii* comprised just 10% of the total cell count but was present for 9 months. *Pseudanabaena limnetica* comprised 22% of total cell count and was present for 13 months. Similarly, *Coelomoron pusillum* (10% cell count, present 13 months) and *Aphanocapsa delicatissima* (8% cell count, present 12 months) presented as fewer cell counts but more consistent presence.

Lake Whangape blue-green algae profile was similar to Waikare and dominated by 3 genera – *Planktolyngbya, Aphanocapsa*, and *Merismopedia* – accounting for 92% of blue-green algae cell counts. These are common to dominant genera found in Waikare. As with Waikare, *Planktolyngbya* was the most dominant genus, accounting for 37% of cell counts, although species distribution was different, with *Planktolyngbya* cf. *tallingii* the predominant species from this genus. For the *Aphanocapsa* genus (36% of cell count) virtually all of the cell count was for the species *A. delicatissima* (36% of cell count). *M. minutissima* (19% of cell count) was the only species identified from the *Merismopedia* genus.

Lake Waahi blue-green algae profile was markedly different to Waikare and Whangape and dominated by 2 genera – cf. *Pannus* and *Microcystis* - accounting for 91% of blue-green algae cell counts. The cf. *Pannus* genus (53% of cell counts) was not identified to species level. *Microcystis* genus (38% of cell counts) was predominantly *M. flos-aquae* (37%), with the other species present in very low cell counts. Despite being the dominant genus in terms of cell counts, cf. *Pannus* was only present for 2 months. The *Microcystis* genus was present for 8 months, although the predominant species – *M. flos-aquae* – was only present for 3 months.

Hamilton Traffic blue-green algae profile was dominated by 3 genera – *Merismopedia, Planktolyngbya*, and *Microcystis* - accounting for 89% of blue-green algae cell counts. The

Merismopedia genus (53% of cell counts) was identified as one species, *M. minutissima*, although was only present for 1 month. *Planktolyngbya* genus (28% of cell counts) was virtually all *P.* cf. *tallingii* 28% of cell counts) but this was only present for 1 month. *Microcystis* genus comprised 9% of the cell count and was present for 6 months. *M. aeruginosa* was the predominant species (4.6%) but was only present for 2 months.

The Elbow blue-green algae profile was dominated by 2 genera – *Pseudanabaena* and *Planktolyngbya* – accounting for 89% of blue-green algae cell counts. *Pseudanabaena* genus (45% of cell counts) was present for 14 months, and although *P. limnetica* was the predominant species (45%), it was present for only 10 months. For the other 4 months the species was not identified. *Planktolyngbya* genus (44%) was present for 3 months with the predominant species being *P. subtilis* (40%).

Table 7. Summary of dominant blue-green algae genus and species by site.

Genus	% cell count	Months present	Species	% cell count	Months present
Waikare					
Planktolyngbya	47	9	*subtilis*	36	6
			cf. *tallingii*	10	9
			galeata	0.003	1
Pseudanabaena	22	14	*limnetica*	22	13
			mucicola	0.2	9
Coelomoron	10	13	*pusillum*	10	13
Aphanocapsa	9	14	*delicatissima*	8	12
			sp.	1	3
			planctonica	0.001	1
Merismopedia	7	7	*minutissima*	6	4
			sp.	1	3
Whangape					
Planktolyngbya	37	9	cf. *tallingii*	31	6
			subtilis	5	4
			sp.	1	3
Aphanocapsa	36	10	*delicatissima*	36	8
			sp.	0.2	2
			cf. *planctonica*	0.01	1
Merismopedia	19	6	*minutissima*	19	6
Waahi					
cf. *Pannus*	53	2	sp.	53	2
Microcystis	38	8	*flos-aquae*	37	3
			aeruginosa	0.4	4
			sp. large, width > 4 µm	0.01	1
			sp. small, width < 4 µm	0.2	4
			blank	0.004	1
Hamilton Traffic					
Merismopedia	52	1	*minutissima*	52	1
Planktolyngbya	28	2	cf. *tallingii*	28	1
			subtilis	0.04	1
Microcystis	9	6	*aeruginosa*	4.6	2
			flos-aquae	0.6	2
			ichthyblabe	0.4	1
			sp. large, width > 4 µm	0.2	1
			sp. small, width < 4 µm	2.1	1
			wesenbergii	1.5	2
Elbow					
Pseudanabaena	45	14	*limnetica*	45	10
			sp.	0.1	4
			galeata	0.02	4
Planktolyngbya	44	3	*subtilis*	40	3
			cf. *tallingii*	4	2

3.4 "Non" blue-green algae genus distribution

A breakdown of dominant "non" blue-green algae into genus level over the sampling period is summarised in Table 8.

Over the sampling period, 48 "non" blue-green algae genera were identified in Waikare (see Appendix 1). *Synedra* (57.1% of total cell count) dominated algae cell counts, and together with the next most abundant genera – *Actinastrum* (12.7% of cell count) and *Synura* (7.0% of cell count) – comprised the majority of total algae cell counts, accounting for nearly 77%. *Synedra* and *Actinastrum* were also prevalent, occurring in 12 and 13 of 14 samples, respectively (Table 8).

Over the sampling period, 61 "non" blue-green algae genera were identified in Whangape (see Appendix 1). The algae profile was different to Waikare, being dominated by *Flagellates/Unicells* (33.7%), which were always present. *Synedra* (9.3%), *Ankistrodesmus* (5.9%) and *Actinastrum* (5.6%) were the next most abundant genus and, together with *Flagellates/Unicells* accounted for over 50% of the total cell counts (Table 8).

Over the sampling period, 47 "non" blue-green algae genera were identified in Waahi (see Appendix 1). The three most abundant genera – namely *Flagellates/Unicells* (15.4%), *Ceratium* (10.9%), and *Oocystis* (10.3%) – only account for around 37% of the total cell counts (Table 8).

River sites had a markedly different "non" blue-green algae genus profile than the lakes.

Over the sampling period, 61 "non" blue-green algae genera were identified in Hamilton Traffic (see Appendix 1). The three most abundant genera – namely *Fragilaria* (36.2%), *Klebsormidium* (20.2%), and *Asterionella* (9.0%) – accounted for 65% of total cell counts and were generally present most of the time (Table 8).

Elbow had a very similar algae profile to Hamilton Traffic. Over the sampling period, 63 "non" blue-green algae genera were identified in Elbow (see Appendix 1). The three most abundant genera – namely *Fragilaria* (27.2%), *Asterionella* (21.1%), and *Klebsormidium* (13.7%) – accounted for 62% of total cell counts and were generally present most of the time (Table 8).

Table 8. Summary of dominant "non" blue-green algae genus by site.

Genus	% cell count	Months present
Waikare		
Synedra	57.1	12
Actinastrum	12.7	13
Synura	7.0	1
Whangape		
Flagellates/Unicells	33.7	14
Synedra	9.3	8
Ankistrodesmus	5.9	9
Actinastrum	5.6	13
Waahi		
Flagellates/Unicells	15.4	13
Ceratium	10.9	6
Oocystis	10.3	13
Hamilton Traffic		
Fragilaria	36.2	14

Genus	% cell count	Months present
Klebsormidium	20.2	13
Asterionella	9.0	11
Elbow		
Fragilaria	27.2	13
Asterionella	21.1	12
Klebsormidium	13.7	13

3.5 Sensor data

3.5.1 PC and Chl-*a* fluorescence

Average and median PC and Chl-*a* sensor fluorescence data and turbidity data for the lake sites were obtained from WRC Hydrotel database and median data summarised in Table 9 and Figure 6. River PC and Chl-*a* sensor fluorescence data and turbidity data[10] for the two river sites are summarised in Table 10 and Figure 7.

Parallel to algae cell counts, PC and Chl-*a* fluorescence generally followed the trend Waikare > Whangape > Waahi > Waikato River sites. At Waahi, the PC sensor recorded very low or negative values for sustained periods of time (May to December 2017), suggesting issues with sensor calibration and stability.

Turbidity was generally variable but markedly higher for the lake sites than the river sites. Elbow site had consistently higher turbidity than Hamilton Traffic site.

[10] Data used are single values – if collection was at the same time as the sensor output – or an average of two values if water collection fell between two sensor values data for 3 outputs.

Table 9. Summary of sensor median PC and Chl-*a* fluorescence and turbidity data for lake sites.

Date	Time	PC	Chl-*a*	Turbidity
Waikare				
24/04/2017	13:35	118 .5	69.7	7.9
24/05/2017	11:15	133.7	72.2	21.9
29/06/2017	13:24	177.7	79.9	19.4
25/07/2017	14:30	192.6	87.0	27.8
24/08/2017	14:45	141.3	74.4	18.9
27/09/2017	10:30	126.6	95.1	26.2
26/10/2017	12:30	202.1	118.5	49.6
4/12/2017	14:00	209.2	45.7	10.2
21/12/2017	13:45	214.1	115.2	46.8
30/01/2018	11:00	82.4	51.0	5.8
22/02/2018	15:30	94.7	85.6	16.6
26/03/2018	12:00	17.6	54.8	5.2
27/04/2018	13:00	48.1	45.8	10.3
25/05/2018	11:30	92.9	79.9	18.2
Whangape				
24/04/2017	11:45	23.4	74.0	7.2
24/05/2017	12:45	7.7	23.9	7.7
29/06/2017	11:55	59.0	53.2	14.3
25/07/2017	15:45	8.9	19.5	16.5
24/08/2017	12:55	43.1	20.2	11.2
27/09/2017	12:15	28.7	52.5	4.8
26/10/2017	11:15	42.6	80.4	60.4
4/12/2017	15:30	52.2	47.5	45.7
21/12/2017	12:10	60.4	69.7	52.1
30/01/2018	12:15	119.9	63.9	131.2
22/02/2018	13:15	76.0	32.9	30.0
26/03/2018	13:15	53.5	16.5	11.7
27/04/2018	11:00	35.3	13.5	22.3
25/05/2018	12:45	22.4	10.3	12.8
Waahi				
24/04/2017	10:01	10.6	10.9	6.2
24/05/2017	13:45	-8.0	5.2	8.6
29/06/2017	10:25	-2.3	11.7	12.6
25/07/2017	16:45	-5.8	4.2	10.9
24/08/2017	10:45	-2.4	7.4	7.7
27/09/2017	13:15	-3.9	4.5	5.4
26/10/2017	10:00	0.5	1.6	13.9
4/12/2017	16:45	-0.5	19.6	8.9
21/12/2017	9:30	-0.6	20.0	21.1
30/01/2018	13:15	11.1	6.5	11.0
22/02/2018	9:30	3.2	11.2	23.6
26/03/2018	14:15	11.1	39.1	12.4
27/04/2018	9:20	39.6	152.9	11.9
25/05/2018	13:45	18.4	91.2	7.7

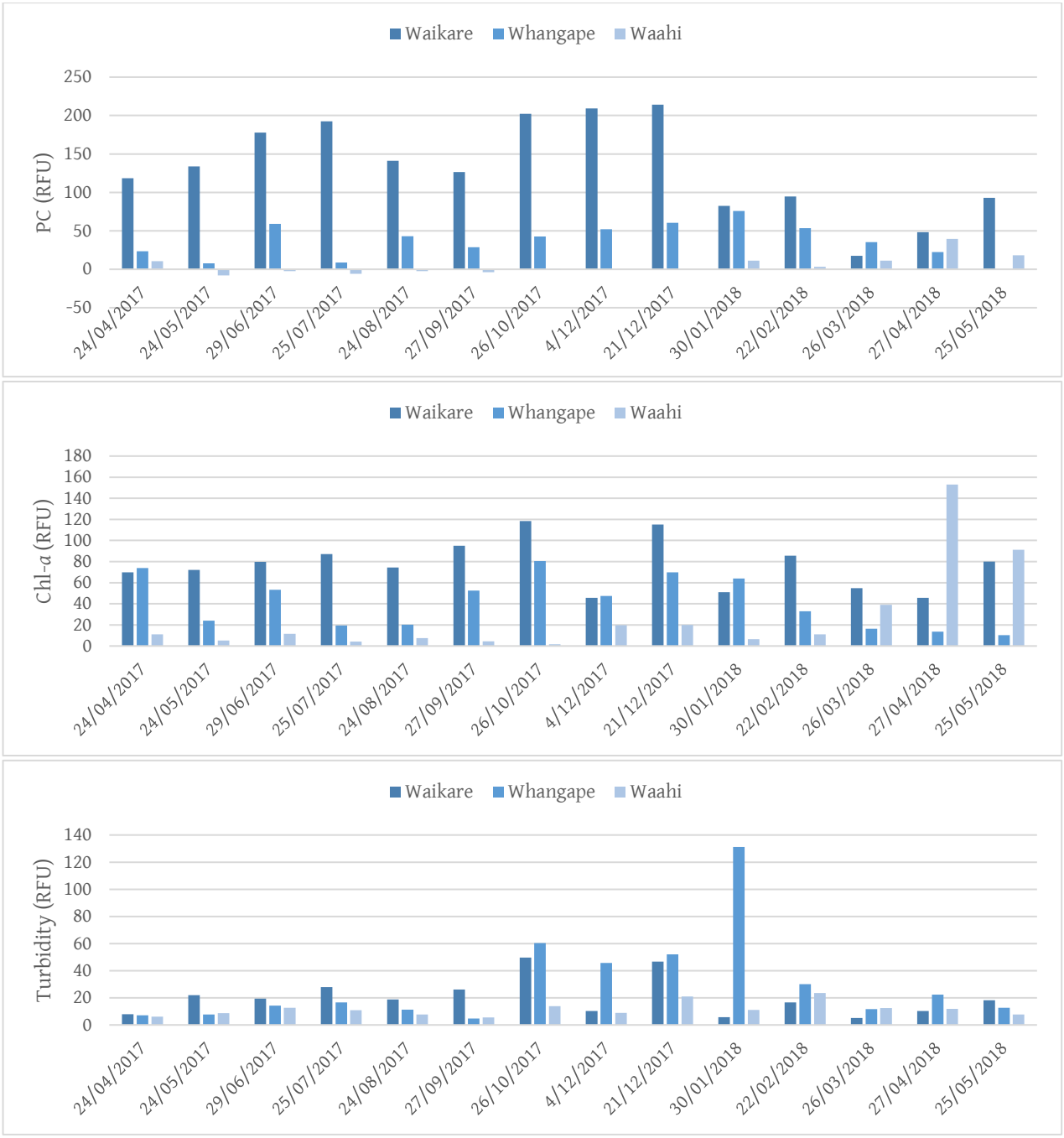


Figure 6. Summary of sensor median PC and Chl-*a* fluorescence and turbidity data for lake sites.

Table 10. Summary of sensor median PC, Chl-*a* fluorescence and turbidity data for river sites.

Date	Time	PC	Chl-*a*	Turbidity
Hamilton Traffic				
1/06/2017	9:35	0.81	0.45	2.5
4/07/2017	10:45	0.14	0.94	3.4
10/08/2017	13:59	0.51	2.42	6.5
5/09/2017	9:26	0.15	5.21	3.4
29/09/2017	10:40	0.27	9.21	7.7
6/11/2017	10:30	0.42	2.42	2.7
18/12/2017	12:30	0.17	4.39	2.5
3/01/2018	8:30	0.13	3.41	2.6
19/02/2018	9:30	0.13	0.19	2.2
26/02/2018	Unknown[1]	0.13	3.07	1.5
4/05/2018	11:41	0.08	0.39	2.0
28/05/2018	11:20	0.70	1.23	2.9
6/06/2018	14:34	0.16	1.62	3.7
19/06/2018	15:50	0.16	1.36	2.8
Waikato River Elbow				
7/06/2017	11:54	0.33	1.73	8.8
3/07/2017	11:05	0.88	1.09	20.6
10/08/2017	10:20	0.15	1.47	14.2
5/09/2017	12:00	0.40	7.61	14.8
28/09/2017	11:00	0.27	15.31	13.4
31/10/2017	10:20	0.62	15.54	21.4
30/11/2017	11:40	0.09	6.26	7.3
20/12/2017	9:10	0.32	13.11	11.8
3/01/2018	11:30	0.37	16.08	5.0
5/02/2018	11:15	0.29	2.92	12.4
26/02/2018	11:18	0.44	7.24	9.5
4/04/2018	11:40	0.24	2.49	3.8
2/05/2018	11:32	0.02	0.08	12.3
6/06/2018	11:59	0.17	4.16	45.9
19/06/2018	11:52	0.12	4.19	16.2

[1] Time of collection was unknown so data average from 9am to 5pm.

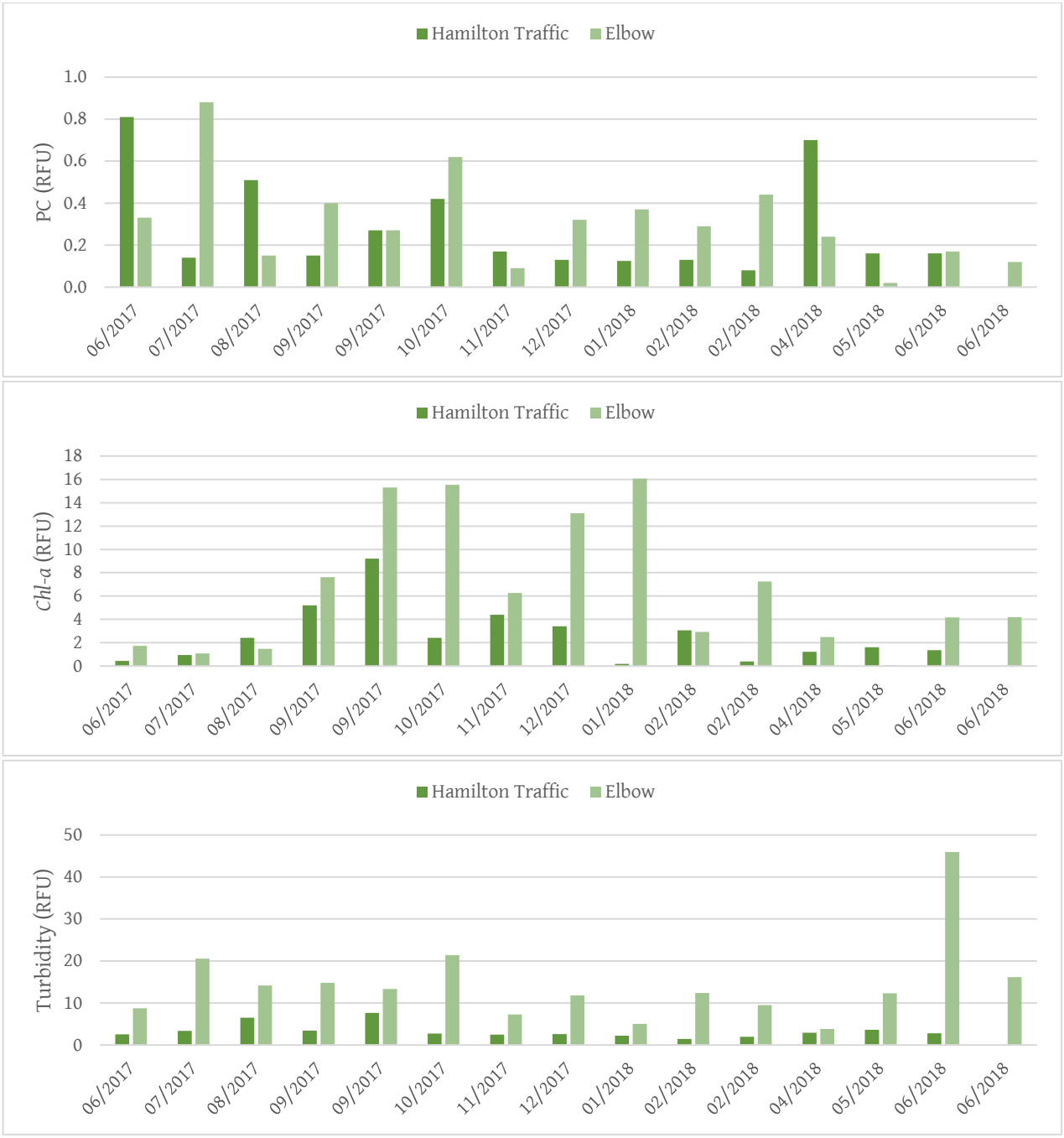


Figure 7. Summary of sensor median PC, Chl-*a* fluorescence and turbidity data for river sites.[11]

3.6 Water measurements

pH and turbidity were measured on site with a hand-held meter, while pH, suspended solids (SS) and Chl-*a* were measured in the laboratory. Water measurements are summarised for lake sites in Table 11 and river sites in Table 12.

[11] As river sites were not sampled always on the same date and at times twice in one month, only month shown.

Table 11. Summary of pH (site and lab), turbidity (site), suspended solids (lab) and Chl-*a* (lab) data for lake sites.

Date	Time	pH (site)	pH (lab)[1]	Turbidity (site) NTU	SS (lab) g/m^3	Chl-*a* (lab) mg/m^3
Waikare						
24/04/2017	13:35	9.0	ND	18.2	27.7	97.9
24/05/2017	11:15	7.0	ND	ND[2]	61.0	142.0
29/06/2017	13:24	8.8	ND	38.1	60.0	78.3
25/07/2017	14:30	8.6	ND	54.5	67.3	119.0
24/08/2017	14:45	9.2	ND	28.2	48.9	202.0
27/09/2017	10:30	8.2	ND	45.7	36.1	149.0
26/10/2017	12:30	8.4	ND	96.0	194.0	260.0
4/12/2017	14:00	10.4	10.1	10.5	27.7	30.2
21/12/2017	13:45	ND[3]	9.0	ND[3]	165.0	147.0
30/01/2018	11:00	8.7	8.8	14.3	41.0	59.7
22/02/2018	15:30	8.9	9.3	26.2	70.7	92.0
26/03/2018	12:00	9.3	9.1	9.8	19.0	23.6
27/04/2018	13:00	8.9	8.6	25.1	37.7	46.0
25/05/2018	11:30	7.9	8.0	41.3	68.3	105.0
Whangape						
24/04/2017	11:45	7.4	ND	20.0	23.7	94.8
24/05/2017	12:45	7.1	ND	ND[2]	17.3	23.0
29/06/2017	11:55	8.6	ND	32.4	40.0	73.4
25/07/2017	15:45	7.7	ND	40.9	39.9	31.4
24/08/2017	12:55	7.6	ND	32.0	26.9	54.6
27/09/2017	12:15	7.7	ND	14.3	10.6	37.5
26/10/2017	11:15	8.1	ND	157.0	182.0	89.2
4/12/2017	15:30	7.9	8.4	70.9	120.0	81.0
21/12/2017	12:10	7.8	8.0	107.0	138.0	60.4
30/01/2018	12:15	7.6	7.7	20.9	36.0	34.4
22/02/2018	13:15	8.0	8.1	87.0	115.0	112.0
26/03/2018	13:15	8.0	8.7	42.9	47.5	28.7
27/04/2018	11:00	7.6	7.8	72.4	87.7	47.0
25/05/2018	12:45	7.7	7.7	29.1	61.6	13.5
Waahi						
24/04/2017	10:01	7.4	ND	15.0	6.3	2.9
24/05/2017	13:45	7.1	ND	ND[2]	8.7	2.0
29/06/2017	10:25	7.8	ND	23.0	17.0	6.1
25/07/2017	16:45	7.7	ND	18.7	15.2	4.1
24/08/2017	10:45	7.6	ND	20.5	13.3	12.0
27/09/2017	13:15	7.6	ND	10.4	5.7	1.9
26/10/2017	10:00	7.7	ND	26.6	54.0	6.5
4/12/2017	16:45	7.5	8.0	9.9	11.9	4.0
21/12/2017	9:30	8.0	8.2	34.0	44.0	13.5
30/01/2018	13:15	8.3	8.3	14.7	22.8	12.3
22/02/2018	9:30	7.9	8.2	41.0	64.5	16.9
26/03/2018	14:15	8.6	8.7	30.9	50.7	36.1
27/04/2018	9:20	8.2	8.6	44.0	63.7	163.0
25/05/2018	13:45	8.2	8.2	25.6	30.2	180.0

ND = no data; [1] pH only measured from December 2017 (see Section 2.2); [2] water quality meter used did not include turbidity (see Section 2.1); [3] data not recorded in WRC database.

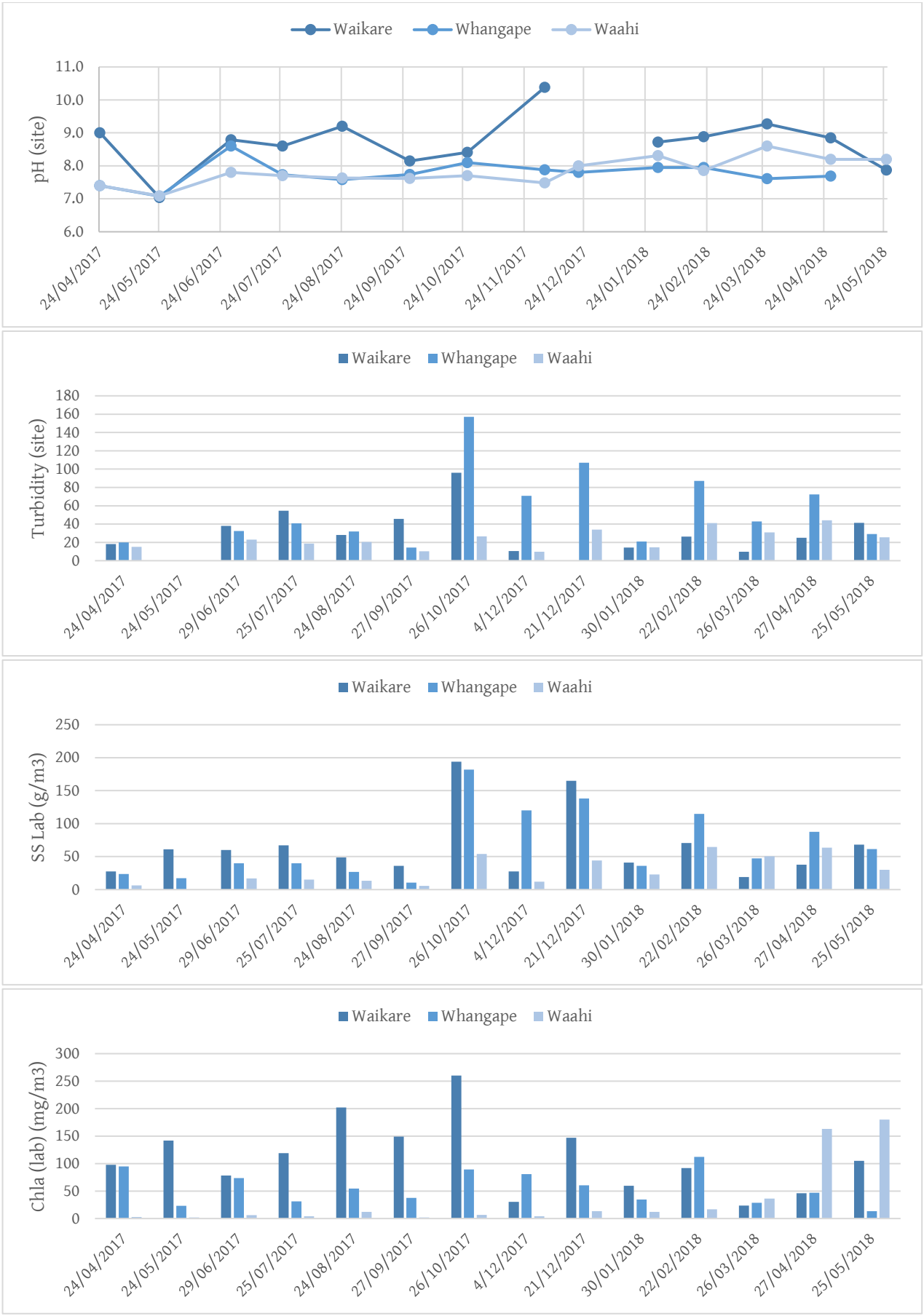


Figure 8. Summary of pH (site), turbidity (site), suspended solids (lab) and Chl-*a* (lab) data for lake sites.

Table 12. Summary of pH (site and lab), turbidity (site), suspended solids (lab) and chl *a* (lab) data for river sites.

Date	Time	pH (site)	pH (lab)	Turbidity (site) NTU	SS (lab) g/m^3	Chl-*a* (lab) mg/m^3
Waikato River Hamilton Traffic						
1/06/2017	9:35	7.4		2.6	4.4	1.3
4/07/2017	10:45	7.4		3.5	6.4	1.7
10/08/2017	13:59	7.4		7.6	10.0	3.6
5/09/2017	9:26	7.4		3.3	6.3	11.9
29/09/2017	10:40	7.4		8.4	13.8	10.0
6/11/2017	10:30	7.5		3.0	6.3	5.3
18/12/2017	12:30	7.7	7.8	2.4	7.8	8.4
3/01/2018	8:30	7.7	ND[1]	2.6	9.5	7.3
19/02/2018	9:30	7.2	ND[2]	3.5	4.1	3.1
26/02/2018	Unknown[3]	ND[3]	7.9	ND[3]	9.8	9.2
4/05/2018	11:41	7.5	7.5	1.7	ND[4]	9.2
28/05/2018	11:20	7.5	7.6	2.8	4.8	3.0
6/06/2018	14:34	7.3	7.6	4.9	6.3	3.4
19/06/2018	15:50	7.5	7.6	3.5	4.7	2.7
Waikato River Elbow						
7/06/2017	11:54	7.3		9.4	13.6	3.8
3/07/2017	11:05	7.3		20.2	26.4	6.6
10/08/2017	10:20	7.2		12.9	19.0	4.7
5/09/2017	12:00	7.4		13.2	23.5	15.6
28/09/2017	11:00	7.3		ND[3]	23.5	14.1
31/10/2017	10:20	7.1		18.5	20.0	36.9
30/11/2017	11:40	6.8		5.1	14.0	6.7
20/12/2017	9:10	7.7	7.9	7.3	16.2	11.1
3/01/2018	11:30	8.0	ND[1]	4.3	4.8	19.6
5/02/2018	11:15	7.3	ND[2]	10.0	21.2	7.2
26/02/2018	11:18	7.4	7.8	9.0	24.4	15.5
4/04/2018	11:40	7.5	7.7	4.5	8.1	8.0
2/05/2018	11:32	7.6	8.0	11.1	13.4	2.2
6/06/2018	11:59	7.1	7.1	63.0	58.7	3.4
19/06/2018	11:52	7.1	7.3	16.0	16.1	4.6

ND = no data; [1] Samples brought to lab when closed over Christmas break. Analysis missed as person not familiar with these samples did the processing. [2] missed by analyst; [3] field data not recorded; [4] filtration volume not recorded by the analyst so no result possible.

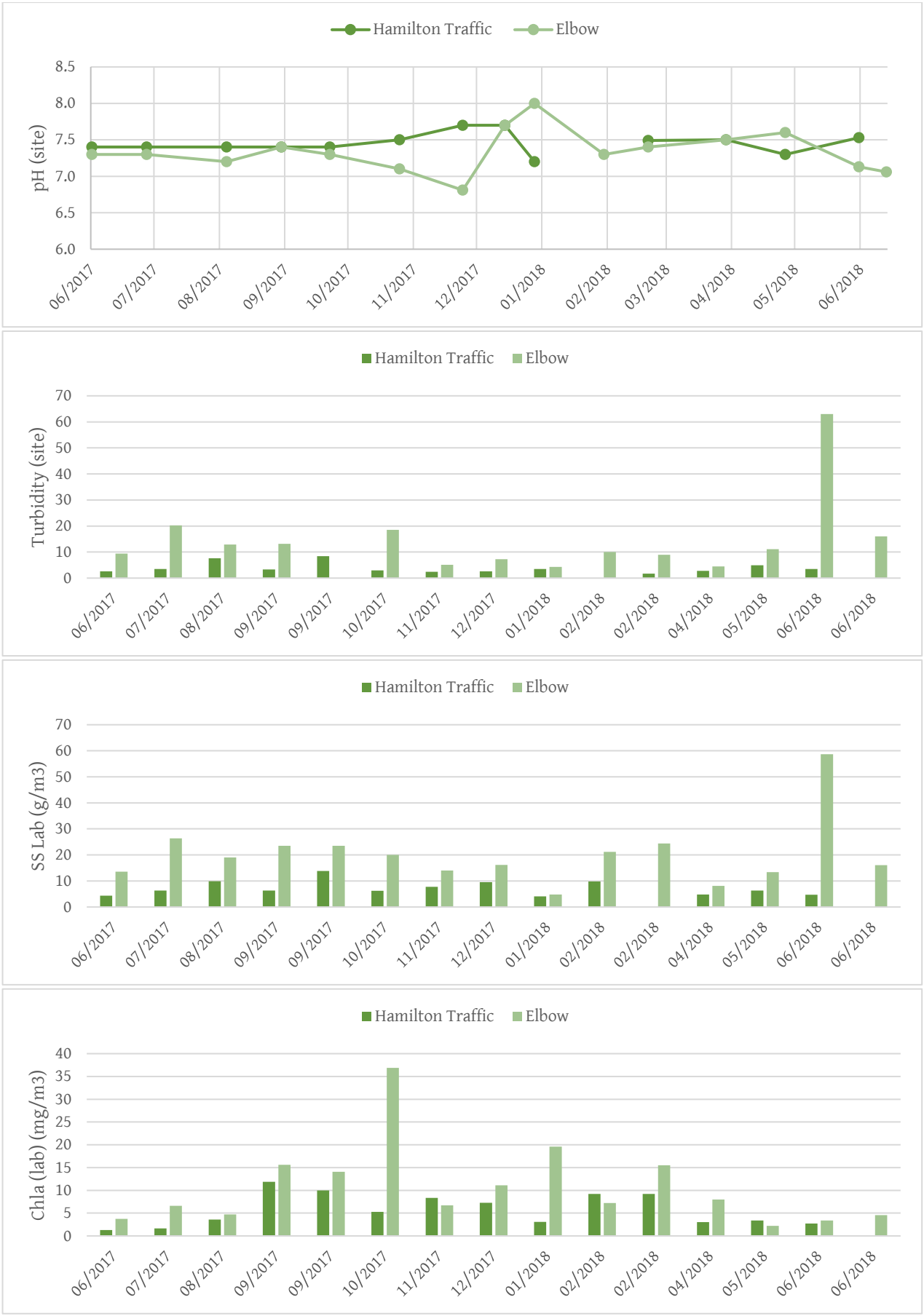


Figure 9. Summary of pH (site), turbidity (site), suspended solids (lab) and Chl-*a* (lab) data for river sites.

4. Regression analyses

4.1 Site-specific regression analysis of sensor PC fluorescence and blue-green algae cell counts

Linear regression analyses were undertaken between PC sensor data and blue-green algae cell counts. These were initially targeted at total algae, but subsequently drilled down to genus and species level for dominant and prevalent algae. Regression results are summarised in Table 13, with graphs presented in Appendix 2.

As is summarised in Table 13, the only significant relationships (P <0.05) between PC sensor fluorescence and blue-green algae cell counts were for total cell count data that had been pooled for all sites (P <0.001) and lake sites only (P <0.001). These pooled data also provided a moderate coefficient of determination with R^2 = 0.57 and 0.51 for all sites and lake sites only, respectively. River site data (pooled or individual) showed very weak relationships with all R^2 < 0.04. Similarly, Waahi provided no relationship (R^2 < 0.01). Waikare shows a weak relationship between PC fluorescence and total blue-green algae cell counts (R^2 =0.18). Whangape showed a moderately weak relationship (R^2 =0.20).

This suggests that n=14[12] for individual sites is not sufficient to provide robust statistical regression analysis for PC sensor fluorescence and total blue-green algae cell counts.

Generally, for individual sites, total blue-green algae cell counts had stronger coefficients of determination than drilling down to dominant or prevalent genus or species. The main exception was Waikare, where for total algae a weak relationship of 0.18 was observed, whereas *Pseudanabaena* – and in particular *P. limnetica* – had a moderately weak relationship of 0.26. Although *Pseudanabaena* was the second most dominant genus present in Waikare (in terms of cell counts) it was always present. This was also a near-significant relationship (P = 0.06) and would likely be strengthened with a larger dataset.

Non-linear regressions (logarithmic, polynomial, power) were also investigated to explain the trends but these showed no consistent improvement to the linear regressions. Furthermore, literature suggests the relationship between PC fluorescence and blue-green algae cell counts is linear. This is discussed in Section 5.

[12] n=13 for Whangape and n=15 for Elbow.

Table 13. Taxonomic breakdown of regression analyses between sensor PC fluorescence and blue-green algae cell counts. Numbers in brackets are % of total cell count and number of samples present.

Site	Taxonomic level			Regression analysis[1]	
	Total	Genus	Species	Coefficient of determination (R^2)	*P*
All sites	Total			0.57	<0.001
Lake sites only	Total			0.51	<0.001
River sites only	Total			0.02	0.52
Waikare	Total			0.18	0.13
		Planktolyngbya (47/9)		0.13	0.20
			P. subtilis (36/6)	0.12	0.23
			P. cf. *tallingii* (10/9)	0.10	0.26
		Pseudanabaena (22/14)		0.26	0.06
			P. limnetica (22/13)	0.26	0.06
		Coelomoron (10/13)		0.01	0.74
			C. pusillum (10/13)	0.01	0.74
		Aphanocapsa (9/14)		0.08	0.33
		Merismopedia (7/7)		0.04	0.49
Whangape	Total			0.20	0.13
		Planktolyngbya (37/9)		0.13	0.23
			P. cf. *tallingii* (31/6)	0.13	0.23
		Aphanocapsa (36/10)		0.17	0.16
			A. delicatissima (36/8)	0.17	0.16
		Merismopedia (19/6)		0.06	0.44
Waahi	Total			<0.01	0.92
		cf. *Pannus* (53/2)		<0.01	1.00
		Microcystis (38/8)		0.01	0.80
			M. flos-aquae (37/3)	0.01	0.80
Hamilton Traffic	Total			0.04	0.47
		Merismopedia (52/1)		0.03	0.52
		Planktolyngbya (28/2)		0.04	0.52
		Microcystis (9/6)		0.11	0.24
Elbow	Total			0.04	0.47
		Pseudanabaena (45/14)		0.16	0.16
		Planktolyngbya (44/3)		0.03	0.55

[1] Significant relationships are underlined (P<0.05).

4.2 Site-specific regression analysis of sensor Chl-*a* fluorescence and "non" blue-green algae cell counts

Linear regression analyses were undertaken between Chl-*a* sensor data and "non" blue-green algae cell counts. These were initially targeted at total algae, but subsequently drilled down to genus level.[13] Regression results are summarised in Table 14, with graphs presented in Appendix 3.

[13] In practice genus taxonomic level was sufficient as there was generally only 1 species for each genus. The exceptions to this were *Flagellates/Unicells* (predominantly <5µm species) and *Oocystis* (predominantly *Oocystis* sp.).

As Table 14 shows – and assessing total counts only – there are considerable inter-site differences in relationships between Chl-*a* fluorescence and total "non" blue-green algae, where lake sites were all weak (R^2 = 0.01 - 0.17) and not significant (P = 0.16 - 0.69), Hamilton Traffic moderately weak (R^2 = 0.28) and significant (P = 0.05), and Elbow moderately strong (R^2 = 0.76) and significant (P <0.001).

In contrast to the PC relationships with blue-green algae cell counts, relationships between Chl-*a* fluorescence and total "non" blue-green algae are not improved by pooling the data. For all sites, only a weak relationship was observed (R^2 = 0.10), although this was significant (P = 0.01). For lake sites only, the relationship was weak (R^2 = 0.07) and not significant (P = 0.10). For river sites only, the relationship was significant (P <0.001) and moderate (R^2 = 0.48).

However, the very weak relationships for the lake sites are partially due to data outliers. Each lake site has a single outlier[14] (see Appendix 3) that when removed increases the strength of the relationships: with Waikare (R^2 = 0.25, P = 0.08) and Whangape (R^2 = 0.29, P = 0.07) increasing to moderately weak, and near-significant and Waahi (R^2 = 0.48, P = 0.01) increasing to moderate and significant. This also impacts the pooled data, with increases to moderately weak and significant for lake sites (R^2 = 0.21, P = 0.004) and all sites (R^2 = 0.25, P <0.001). River site data had no obvious data outliers.

Drilling down to genus provided a few insights into the weak relationships. This was consistent with blue-green algae and PC relationships (Section 4.1).

The data outliers observed with total cell counts at lake sites mentioned above were traced to individual genus. The Waikare outlier (August 2017) was due to an unusually high cell count for *Synedra* (1,227,876 cells/mL), which was the dominant genus present in Waikare with 57% of the total cell count. However, although *Synedra* was present for 12 of 14 months, the majority (96%) of the cell counts was from the August 2017 sampling. Furthermore, this count accounted for 55% of total cell counts from Waikare over 14 months. A similar phenomenon was seen with Whangape and Waahi, however the data outliers were due to a single high cell count for flagellates/unicells at each site.

This highlights how a single data outlier for a relatively low number of datapoints (14) can significantly skew the data.

In summary, "non" blue-green algae relationships with Chl-*a* sensor fluorescence show a somewhat reversed situation to blue-green algae with PC sensor fluorescence. River sites generally have stronger and more significant relationships than lake sites, drilling down to genus shows improvement for some sites/genera, whereas pooling total data does not improve the situation.

[14] Outliers are: Waikare total "non" blue-green algae 1,273,403 cells/mL on 24/08/17; Whangape total "non" blue-green algae 269,053 cells/mL on 21/12/17; Waahi total "non" blue-green algae 9,513 cells/mL on 21/12/17.

Table 14. Taxonomic breakdown of regression analyses between sensor Chl-*a* fluorescence and "non" blue-green algae cell counts.

Site	Taxonomic level[1]		Regression analysis[2]	
	Total	Genus	Coefficient of determination (R^2)	*P*
All sites	Total		0.10	0.01
Lake sites only	Total		0.07	0.10
River sites only			0.48	<0.001
Waikare	Total		0.01	0.69
		Synedra (57/12)	0.001	0.93
		Actinastrum (13/13)	0.19	0.12
		Synura (7/1)	0.23	0.08
Whangape	Total		0.17	0.16
		Flagellates/Unicells (34/14)	0.14	0.21
		Synedra (9/8)	0.13	0.22
		Ankistrodesmus (6/9)	0.18	0.15
		Actinastrum (6/13)	0.37	0.03
Waahi	Total		0.06	0.38
		Flagellates/Unicells (15/13)	0.003	0.86
		Ceratium (11/6)	0.91	<0.001
		Oocystis (10/13)	0.01	0.76
Hamilton Traffic	Total		0.28	0.05
		Fragilaria (36/14)	0.04	0.51
		Klebsormidium (20/13)	0.26	0.06
		Asterionella (9/11)	0.21	0.10
Elbow	Total		0.76	<0.001
		Fragilaria (36/14)	0.30	0.03
		Asterionella (36/14)	0.25	0.06
		Klebsormidium (36/14)	0.30	0.03

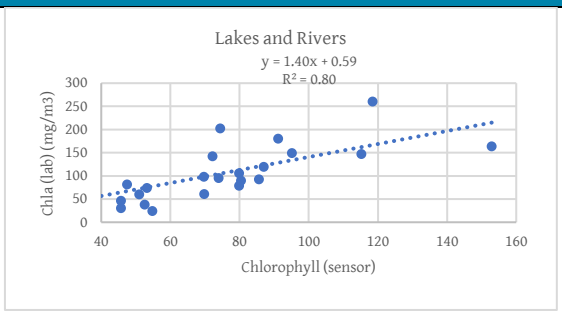
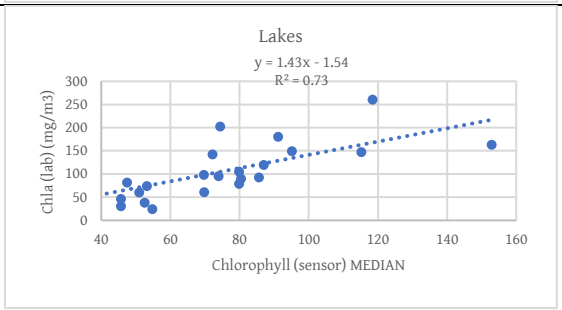
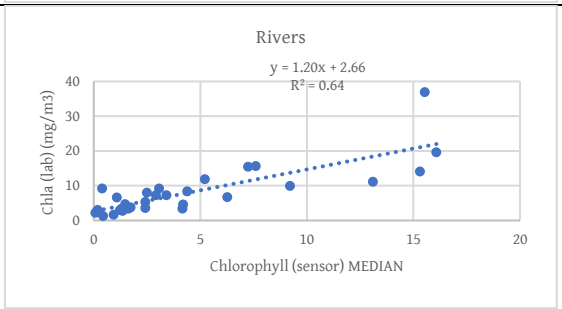
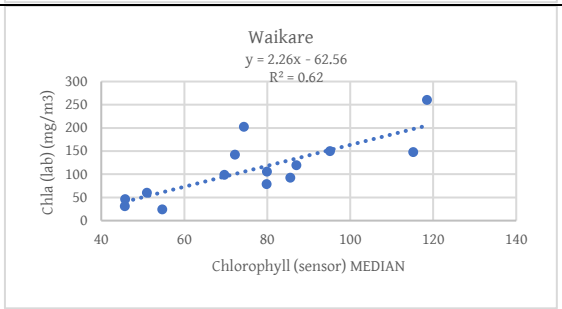
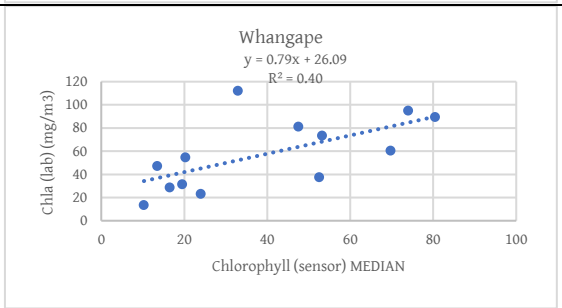
[1] Numbers in brackets are % of total cell count and number of samples present.
[2] Significant relationships are underlined (P<0.05).

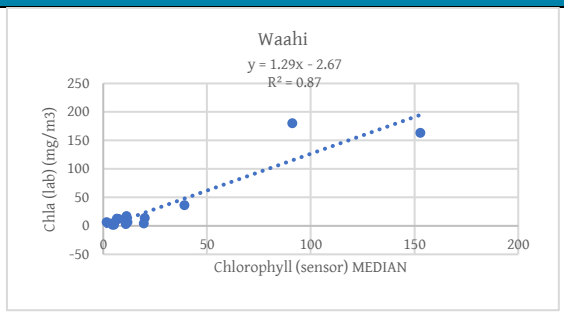
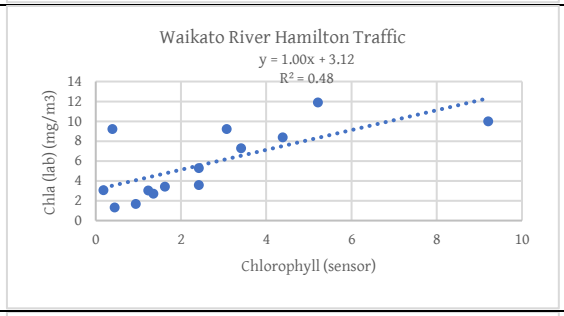
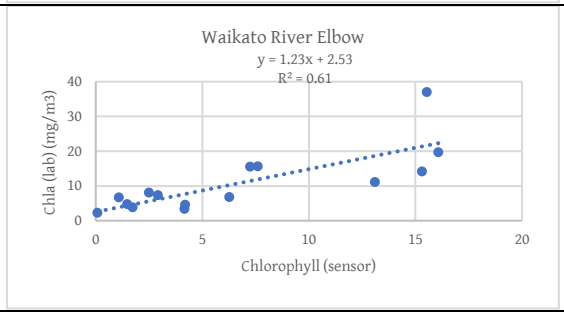
4.3 Site-specific regression analysis of sensor Chl-*a* fluorescence and water Chl-*a* concentration

The relationship between sensor Chl-*a* fluorescence and Chl-*a* concentration measured by the laboratory from water samples collected is summarised in Table 15.

The relationship appears to be site-specific and ranges from moderate to strong. When all site data are pooled, a strong R^2 of 0.80 is obtained, which is only slightly lower (moderately strong) for lake sites only (R^2 = 0.73) or river sites only (R^2 = 0.64). For individual lakes there is some variation. Waikare shows a moderately strong relationship (R^2 = 0.62), Whangape moderate (R^2 = 0.40), and Waahi strong (R^2 = 0.87). The Waahi relationship is driven – to some extent – by the two high values, which when removed affords a weaker – but still moderately strong – relationship (R^2 = 0.62). The river sites exhibit moderate to moderately strong relationships (Hamilton Traffic R^2 = 0.48 and Elbow R^2 = 0.61). All relationships are significant (*P* <0.05).

Table 15. Regression analysis summary for sensor Chl-*a* fluorescence and water Chl-*a* concentration.

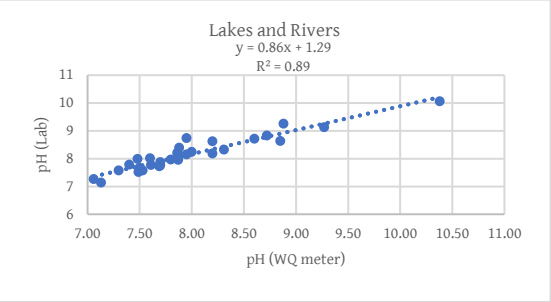
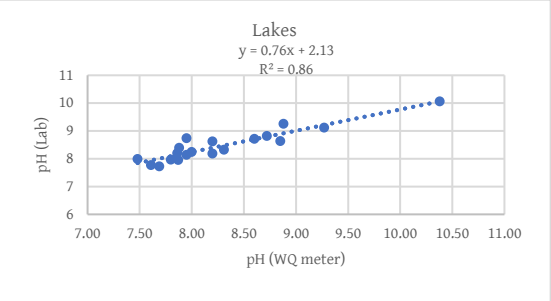
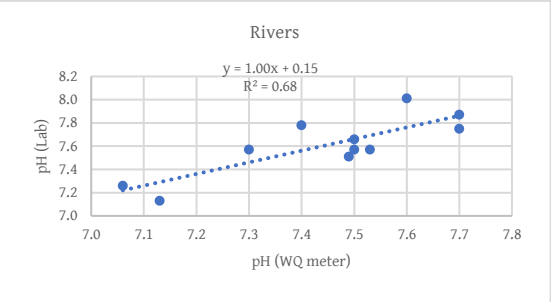
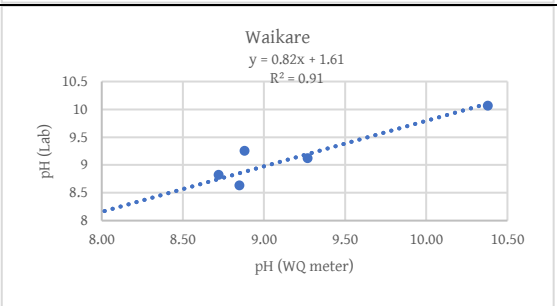
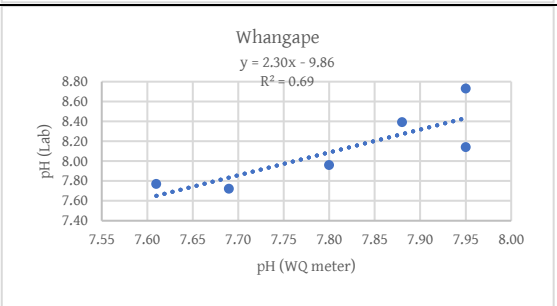
Site	N	Coefficient of determination (R^2)	*P*	Graph
All sites	71	0.80	<0.001	
Lake sites	42	0.73	<0.001	
River sites	29	0.64	<0.001	
Waikare	14	0.62	0.001	
Whangape	14	0.40	0.019	

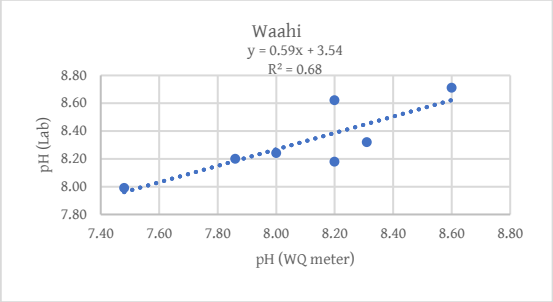
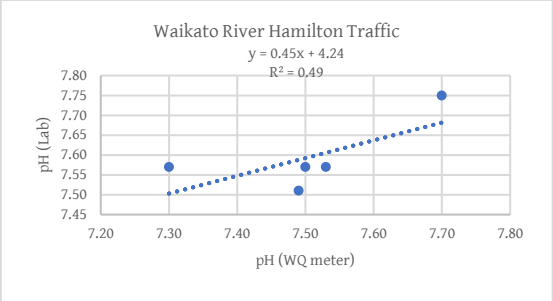
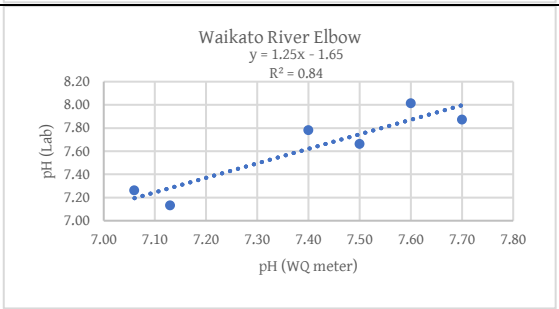
Site	N	Coefficient of determination (R^2)	*P*	Graph
Waahi	14	0.87	<0.001	
Hamilton Traffic	14	0.48	0.006	
Elbow	15	0.61	0.001	

4.4 Site-specific regression analysis of laboratory pH measurement and pH from on-site hand-held meter

The relationship between laboratory pH measurement and on-site pH measurement is summarised in Table 16. Despite the low number of datapoints (n=6 or 7) for each site, relationships range from moderate (Hamilton Traffic R^2 = 0.49) to strong (Waikare R^2 = 0.91) for individual sites. Despite the low number of datapoints, all relationships are significant (*P* <0.05) except for Hamilton Traffic (*P* 0.19). Pooling the data creates moderately strong relationships (R^2 = 0.77) for river sites only and strong relationships for all sites (R^2 = 0.89) and for lake sites only (R^2 = 0.86) over a pH range of 7 to 10.5.

Table 16. Regression analysis summary for laboratory pH measurement and pH from on-site hand-held meter.

Site	N	Coefficient of determination (R^2)	*P*	Graph
All sites	33	0.89	<0.001	
Lake sites	21	0.86	<0.001	
River sites	12	0.77	<0.001	
Waikare	7	0.91	0.003	
Whangape	7	0.69	0.040	

Site	N	Coefficient of determination (R^2)	*P*	Graph
Waahi	7	0.68	0.022	
Hamilton Traffic	5	0.49	0.190	
Elbow	6	0.84	0.010	

4.5 Site-specific regression of sensor turbidity and laboratory suspended solids

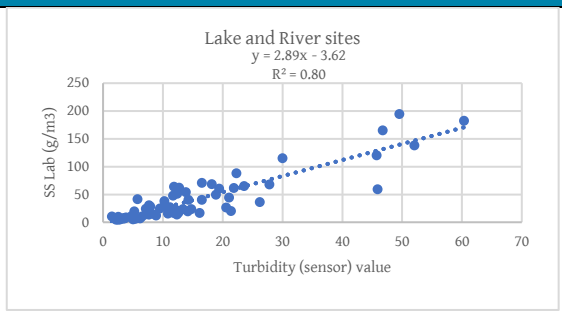
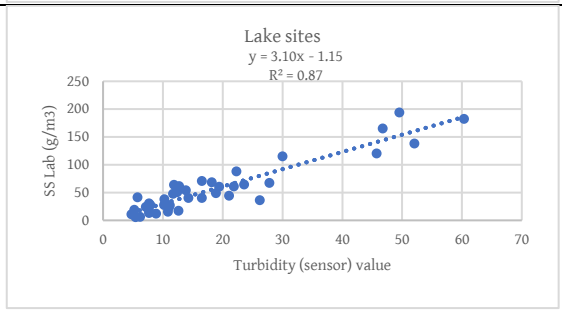
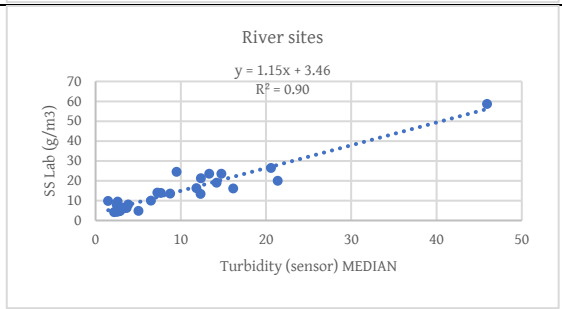
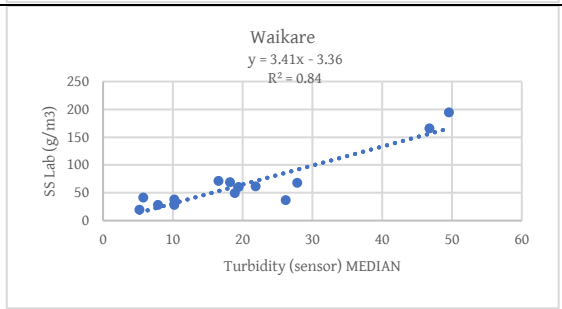
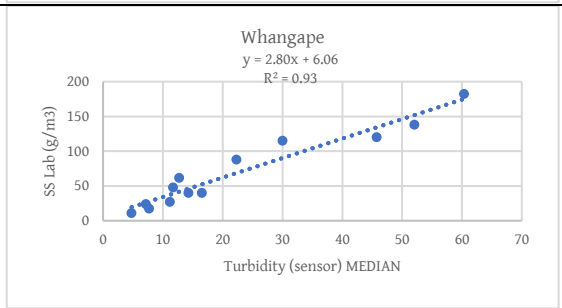
The relationship between laboratory suspended solids measurement and sensor turbidity is summarised in Table 17.

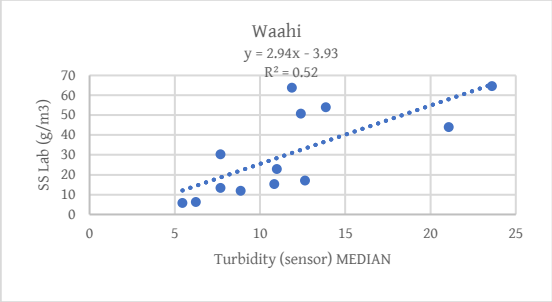
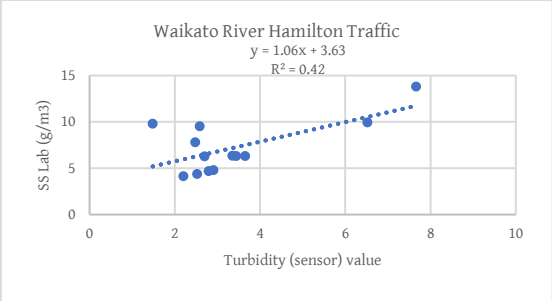
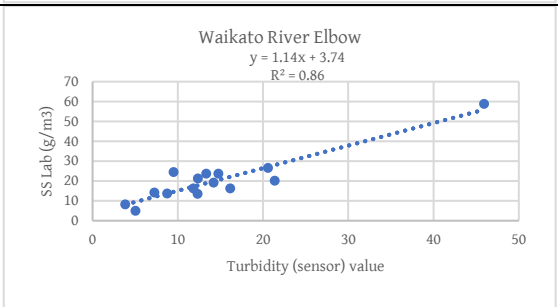
The strength of the relationship is variable for each site ranging from moderate (R^2 = 0.42) at Hamilton Traffic to strong (R^2 = 0.93) at Whangape. All relationships are significant (*P* <0.05).

As the turbidity and suspended solids range increases, the strength of the relationship appears to increase. For example, Hamilton Traffic has a moderate relationship (R^2 = 0.42) and sensor turbidity ranges from 1.5 to 7.7 NTU, while suspended solids ranges from 4.1 to 13.8 g/m^3. In contrast, Waikare has a strong relationship (R^2 = 0.84) and sensor turbidity ranges from 5.2 to 49.6 NTU, while suspended solids ranges from 19 to 194 g/m^3.

When data are pooled, strong relationships are observed, with R^2 = 0.80 for all sites, R^2 = 0.87 for lake sites and R^2 = 0.90 for river sites.

Table 17. Regression analysis summary for sensor turbidity and laboratory suspended solids.

Site	N	Coefficient of determination (R^2)	*P*	Graph
All sites	59	0.80	<0.001	
Lake sites	41	0.87	<0.001	
River sites	28	0.90	<0.001	
Waikare	14	0.84	<0.001	
Whangape	14	0.93	<0.001	

Site	N	Coefficient of determination (R^2)	*P*	Graph
Waahi	13	0.52	0.005	
Hamilton Traffic	13	0.42	0.016	
Elbow	15	0.86	<0.001	

The relationship of turbidity (sensor) with turbidity (WQ meter) was also explored to assess reliability of the turbidity sensor at each site. Results were:

- All sites: Strong and significant relationship (R^2 = 0.84, *P* <0.001);
- Lake sites: Strong and significant relationship (R^2 = 0.87, *P* <0.001);
- River sites: Strong and significant relationship (R^2 = 0.94, *P* <0.001);
- Waikare: Strong and significant relationship (R^2 = 0.96, *P* <0.001);
- Whangape: Strong and significant relationship (R^2 = 0.87, *P* <0.001);
- Waahi: Moderate and significant relationship (R^2 = 0.51, *P* 0.006);
- Hamilton Traffic: Strong and significant relationship (R^2 = 0.94, *P* <0.001);
- Elbow: Strong and significant relationship (R^2 = 0.96, *P* <0.001).

These data show that the turbidity sensor is generally strongly and significantly correlated with turbidity water quality measurements. The only exception is Waahi, where the relationship was moderate (R^2 = 0.51). As shown in Appendix 4, the data for Waahi appears more variable than the other sites. Removal of 1 outlier (Turbidity sensor 11.87 NTU, Turbidity WQ meter 44.0 NTU) strengthens this relationship (R^2 = 0.77).

Regression analysis summary

Coefficient of determination (R^2) for biotic indices (blue-green and "non" blue-green algae cell counts) with applicable sensor fluorescence data (phycocyanin and Chl-*a*, respectively) were generally weak, unless the data were pooled to increase the statistical power of the regression analysis. Conversely, Chl-*a* sensor fluorescence and abiotic indices (hand-held water quality measurement of pH and turbidity) generally had markedly stronger relationships with appropriate laboratory measurements (Chl-*a*, pH, SS, respectively), despite lower number of datapoints.

Pooling of data generally improved the strength and significance of relationships, suggesting greater statistical robustness requires more datapoints than generally obtained for individual sites ($n<15$), at least for biotic indices. However, this is not always the case and confounding factors are discussed in Section 5.2. Data outliers had a marked effect on strength and significance of some relationships, especially for Chl-*a* sensor fluorescence relationship with "non" blue-green algae cell counts.

Non-linear regressions (logarithmic, polynomial, power) were also briefly investigated but these showed no consistent improvement to linear regressions. Furthermore, they are inconsistent with linear relationships reported in the literature (see Section 5.1).

For algae regressions, drilling down to genus and/or species taxonomic level generally reduced the strength of R^2. This was partly due to different dominant genus/species between sites and sporadic occurrence of many genus/species over the 14-month timeframe.

5. Discussion

The strength of regression analyses for each variable assessed was varied, particularly for algae and sensor fluorescence. An understanding of the complexities involved in these relationships is summarised below from literature.

5.1 Literature review on PC sensor fluorescence as a tool for cyanobacteria monitoring

Within New Zealand, a recent paper by Hodges et al (2018) assessed 5 different PC sensors to understand variability, specifically around temperature and algae morphology. The TriLux (used here for lake sites) and YSI sensors (used here for river sites) were the same as used in the current study, although TriLux was not assessed in field trials, so comparisons to the current study are limited. The main difference to the current study was that Hodges et al (2018) compared PC sensor fluorescence in the field (Lake Rotorua, Kaikoura, a small shallow eutrophic lake dominated by cyanobacteria) against biovolume and laboratory PC fluorescence from filtered cells. Furthermore, sensor data were converted from raw RFU readings to PC concentration, based on calibration curves from a C-phycocyanin standard. Relationships were determined using linear regression. These showed a strong and significant positive relationship between laboratory PC fluorescence and sensor PC fluorescence ($R^2 > 0.7$, n=29, $P < 0.001$), but weak relationships between sensor PC fluorescence and biovolume for 2 sensors ($R^2 = 0.22 – 0.29$, n=29, $P < 0.001$) and a non-significant relationship for the third, YSI sensor ($R^2 = 0.29$, n=29, $P > 0.4$).

Of the 5 sensors, TriLux and YSI had the lowest upper range of PC concentrations (~1,200 µg/L and ~2,400 µg/L, respectively cf. ≥12,000 µg/L for the other sensors) and did not produce a linear response above these values. Furthermore, the YSI sensor had low short-term precision.

Internationally, literature related to cyanobacteria appear to centre around assessing their human health risks in drinking water reservoirs, especially real-time changes in cyanobacteria levels, for which PC sensors are being assessed.

Izydorczyk et al (2005) assessed the relationship between PC fluorescence and cyanobacterial biomass (mg freshweight per L) during *Microcystis aeruginosa* blooms in a drinking water reservoir. With a sample size of 32 they noted a positive and significant linear relationship ($r = 0.65$, $P < 0.05$)[15] when cyanobacterial biomass was below 15 mg/L and the dominant species was *Microcystis aeruginosa*. Their threshold of 15 mg/L for intensive blooms equates to approximately 80,000 cells/mL (based on *Microcystis aeruginosa* as the dominant species). A positive PC fluorescence was observed above 0.2 mg/L (around 1,100 cells/mL), which was an effective threshold of detection.

Brient et al (2008) assessed the relationship between PC sensor fluorescence and cyanobacterial biomass (cell count and biovolume) on nearly 800 samples from 35 waterbodies in Western France, of which many were used for recreation and some for drinking water. The water bodies were similar to those in this study in that they are all eutrophic, weakly mineralised (< 350 µS/cm), relatively turbid (Secchi depth < 1.5m) and shallow (75% less than 3 m). Furthermore, the cyanobacteria composition was very heterogeneous.

[15] Izydorczyk et al (2005) and McQuaid et al (2011) reported R not R^2. For consistency, R was converted to R^2 for comparison (see Table 18).

Brient et al (2008) also set "thresholds" for their PC sensor. They stated - with a culture of *Planktothrix agardhii* – the probe was able to measure down to 1700 cells/mL. They noted saturation of the fluorescence signal during bloom conditions, potentially leading to non-linear relationships, and although they did not state the sensor reading where this occurs, they noted that a shutter could be used to reduce fluorescence under bloom conditions.

Despite the highly heterogeneous phytoplankton profile, large number of sites and evidence for external effects, Brient et al (2008) established a moderately strong linear relationship ($R^2 = 0.73$) between sensor PC fluorescence and cyanobacterial cell counts. This suggests strongly that a large dataset will provide enough statistical power to minimise external effects and provide a decent calibration curve for blue-green algae cell counts from PC sensor fluorescence.

A group from Quebec, Canada, reported two separate studies (McQuaid et al., 2011; Zamyadi et al., 2012b) assessing PC sensor fluorescence to approximate cyanobacterial abundance in drinking water reservoirs. Both studies noted eutrophic and poor water conditions and shallow reservoirs (2-6 m depth).

McQuaid et al (2011) stated that their sensor (YSI 6600) was unable to properly estimate the PC concentration above 100 RFU, and they removed any data above this from statistical analyses. They also removed negative values, which represented readings below the probe's lower limit of detection. They stated a significant linear relationship ($r = 0.68$, $n = 26$, $P < 0.01$) between PC fluorescence and cyanobacterial biovolume but did not comment on the effect that removal of high and negative RFU values had on the strength or significance of the regressions.

Zamyadi et al (2012b) used the same sensor as McQuaid et al (2011) to measure PC, Chl-*a*, pH, DO, conductivity, temperature and turbidity in parallel. They noted a moderately strong and significant linear relationship between PC fluorescence and extracted PC ($R^2 = 0.79$, $n = 33$, $P < 0.01$), while a moderately weak but significant relationship between Chl-*a* fluorescence and extracted Chl-*a* ($R^2 = 0.23$, $n = 33$, $P < 0.01$). The authors rationalised that this was because the Chl-*a* probe could detect the majority of eukaryotic algal Chl-*a* but only a small fraction of cyanobacterial Chl-*a*, and cyanobacteria were dominating the algae profile.

Kong et al (2014) used laboratory and field studies to investigate relationships between cyanobacterial abundance, biovolume, cylindrospermopsin concentration and PC fluorescence in a freshwater reservoir in Macau that experiences cyanobacterial blooms. With a sample size of 50 they noted a moderately strong linear relationship ($R^2 = 0.71$) between PC sensor fluorescence and cyanobacteria cell numbers and a slightly stronger linear relationship with biovolume ($R^2 = 0.77$).

Kong et al (2014) separated data into a period of low cyanobacteria cell counts which they defined as <40,000 cells/mL, and a period of high cyanobacteria cell counts (>120,000 cells/mL). For the period of low cell counts ($n = 12$) they found a stronger relationship ($R^2 = 0.90$ for both cell counts and biovolumes) and attributed this to the relatively constant phytoplanktonic structure and composition under stable conditions, where cyanobacteria dominated (86–99% cell counts), among which *Psedanabaena* was dominant, comprising 82–99% of cyanobacteria cell counts, followed by *Cylindrospermopsis* and *Dactylococcopsis*, both with <10% of cyanobacteria cell counts. Significantly, "non" blue-green algae counts (*Chlorophyta* and *Bacillariophyta*) were insignificant.

Linear relationships are consistently observed in the literature between PC sensor fluorescence and a co-variable surrogate for PC – such as cyanobacteria cell count, biovolume, biomass, and extracted PC (Table 18).

Table 18. Summary of literature PC sensor relationship (in the field) with various co-variables.

Coefficient of determination (R^2)	N	Co-variable	Unit	Reference
<0.30	29	Biovolume	mm^3/L	Hodges et al (2018)
0.42	32	Biomass	mg/L	Izydorczyk et al (2005)
0.46	26	Biovolume	mm^3/L	McQuaid et al (2011)
0.51	42	Cell count	cells/mL	Current study – lake sites
0.57	72	Cell count	cells/mL	Current study – all sites
>0.7	29	Extracted PC	µg/L	Hodges et al (2018)
0.71	50	Cell count	cells/mL	Kong et al (2014)
0.73	800	Cell count	cells/mL	Brient et al (2008)
0.77	50	Biovolume	mm^3/L	Kong et al (2014)
0.79	33	Extracted PC	µg/L	Zamyadi et al (2012)

As is summarised in Table 18, the relationship is positive, but the strength varies substantially from <0.30 to 0.79. As is shown in the current study, the number of datapoints influences the strength of the relationship, however there appears to be an effect of diminishing returns. For example, when all the data are pooled (n=72), a moderate coefficient of determination (R^2 = 0.57) is observed, which is only marginally weaker for lake sites only (R^2 = 0.51, n=42). This is supported by literature relationships of PC with cyanobacteria cell counts (i.e. comparable to the current study), where there was only a marginal strengthening of the relationship observed between Kong et al (2014) (R^2 = 0.71) and Brient et al (2008) (R^2 = 0.73) despite the number of datapoints increasing from 50 (Kong et al., 2014) to 800 (Brient et al., 2008).

Literature relationships between PC fluorescence and biovolume are variable. Hodges et al (2018) showed a moderately weak relationship (R^2 <0.30) with n=29, McQuaid et al (2011) a moderate relationship (R^2 =0.46) with n=26 and Kong et al (2014) a moderately strong relationship (R^2 =0.77) with n=50. Although based on limited studies, this suggests confounding factors are significant here.

The strongest consistent relationship was between PC fluorescence and extracted PC. Hodges et al (2018) showed a moderately strong relationship (R^2 >0.70) with n=29, with a similar relationship (R^2 =0.79) noted by Zamyadi et al (2012) with n=33. This suggests that measurement of extracted PC may reduce some of the variability involved when enumeration of blue-green algae and "non" blue-green algae is undertaken.

5.2 Factors influencing phycocyanin sensor effectiveness

The effectiveness of the phycocyanin sensor as a surrogate for lab-derived phycocyanin concentrations or cyanobacterial measures such as biovolume or cell counts may be influenced by a range of biotic and abiotic factors.

Hodges et al (2018) noted that large colonial aggregates caused spikes in their sensor results, suggesting that sample heterogeneity was the main cause of the poor relationship observed in their study.

Izydorczyk et al (2005) noted that minimal PC fluorescence was observed when the phytoplankton community was dominated by diatoms and green algae, even if the total phytoplankton community biomass was high, hinting that these species may contribute a minimal amount of PC fluorescence, which is supported by subsequent studies (e.g. Brient et al., 2008; McQuaid et al., 2011).

Brient et al (2008) noted a number of factors that could have varying effects on PC sensor fluorescence measurement. While natural light had no effect, turbidity reduced the PC signal which was due primarily to the number of particles and their size. These authors also noted that false PC sensor positives were observed when large numbers (100,000 cells/mL) of chlorophyta (green algae) *Scenedesmus opoliensis* and the diatom *Asterionella formosa* were present.

In the current study we noted a significant negative relationship between turbidity and PC sensor measurement at Lake Waikare but not at other lakes (data not shown).

McQuaid et al (2011) suggested that variability in PC fluorescence and cyanobacterial biovolume could be partially due to experimental bias (errors of instrumentation, sample preservation, and counting technique). Genetic variability in the production of PC within cyanobacteria species has also been demonstrated (Bañares-España et al., 2007). Also noted were false positives for chlorophyceae (green algae) and bacillariophyceae (diatoms) when high proportion of the biovolume.[16] They also noted that less important factors could cause variability. For example, the amount of PC produced is dependent on the life stage; fluorescence in surface water samples may be affected by pre-saturation of PC pigments by light; turbidity can interfere with sensor readings.

Changes in cyanobacteria profile is site and season specific, and this variation could go some way to explaining the differences in relationships observed between sites in our study (see Appendix 1).

Although Kong et al (2014) did not directly investigate the sources of interference and bias in their study, they referenced a PhD thesis (Zamyadi, 2011) that showed that the sources of interference in probe readings includes fluctuation in the PC concentrations per cell, specificity of the light source used in the probe, and abiotic factors, particularly water turbidity. The pertinent information from the thesis was published (Zamyadi et al., 2012a), which revealed the turbidity analyses were undertaken in the laboratory. Nevertheless, their results showed turbidity interfered with PC fluorescence, supporting the findings from Brient et al (2008). Following this, Kong et al (2014) surmised that the much lower relationship between PC and cyanobacteria cell number in environmental samples than in pure cultures observed in their study confirms the interferences of the environmental and/or biotic factors, such as cyanobacteria species composition and turbidity.

[16] We note high proportion of "non" blue-green algae occurs in the current study in Whangape over cooler months, Waahi all months except summer, Hamilton Traffic all months, with Elbow all months expect winter.

Operational causes of variability include saturation of the sensor in periods of cyanobacterial bloom leading to non-linear relationships. The level of saturation is also sensor specific. At the other end of the scale, negative readings – presumably as a result of poor calibration and/or fouling of the sensor – can affect variability. In the current study, this was observed in Waahi for half the sensor readings (n=7). For the current study, removal of negatives values from regression analysis of PC fluorescence with total blue-green algae cell counts for Waahi produced a non-significant (P = 0.46) and negative weak relationship (R^2 = 0.11).

Discussion summary

The linear regressions between PC fluorescence and blue-green algae cell counts noted in the current study are consistent with literature. Comparisons of the strength of these relationships between the current study and literature are confounded due to complexities involved. These include:

- Different methodologies;
    - blue-green algae enumeration by cell count and/or biovolume, or PC fluorescence from extracted blue-green algae;
    - different sensor precision and sensitivity;
    - number of datapoints;
    - use (or not) of thresholds to deal with "upper" and "lower" bound sensor fluorescence issues;
    - experimental bias.
- Varying external effects from biotic and abiotic factors;
    - Different site and season specific algae profile;
        - Species-specific PC fluorescence;
        - False-positives from "non" blue-green algae (especially diatoms and green algae) when blue-green algae are present in low proportion of total algae;
    - Genetic and life stage variability of PC production in blue-green algae;
    - Algae heterogeneity (aggregation);
    - Varying and lesser known abiotic effects, specifically turbidity, temperature, and light.

6. Recommendations

This study has set the foundation for complementing water sampling and enumeration of blue-green algae cell counts and/or biovolume with phycocyanin sensor fluorescence. The ultimate goal will be to have strong site-specific relationships at all lakes and rivers of concern in the Waikato region and beyond.

However, there needs to be more information gathered around biotic and abiotic interferences, and their impact at each site. Also, a greater understanding of blue-green algae and "non" blue-green algae species dynamics is necessary for each site. This would need to include consistency of any seasonal variation and potential spatial variation at lake sites and different river reaches. More information is needed on sensor capabilities and accuracy.

Setting upper and lower sensor thresholds may be necessary to improve the variability. However, there are site and sensor specific logistical issues with this. For example, Waikare sensor readings are consistently over 100 RFU (9 out of 14 sampling events), suggested as the upper limit for the YSI probe (McQuaid et al., 2011). The TriLux probe used in this study for lake monitoring has an upper range 50% that of the YSI probe (Table 1 in Hodges et al., 2018), suggesting that a more realistic upper limit for the TriLux probe should be 50 RFU. This would make the TriLux probe – based on the current study data – unsuitable most of the time (12 out of 14 sampling events) for Waikare, approximately half the time for Whangape (6 out of 14 sampling events), but suitable for Waahi all the time. For Waikare and Whangape, sensors described in Hodges et al (2018) – namely from Manta and Turner – should provide a much higher upper threshold (10x) than TriLux and should be considered for sites where frequent cyanobacteria blooms occur, i.e. Waikare and Whangape.

Therefore, recommendations for the future include:

- Extending the dataset at the current sites to increase the number of datapoints and statistical power. This will also allow for assessment of site-specific stability of algae species throughout repeated seasons;
- At critical times of the year – specifically in times of high and low blue-green algae cell counts – increase the number of water monitoring events;
    - For high cell counts – and restricted primarily to Waikare and Whangape where blooms are frequent and of high intensity – this would be useful for assessing a different sensor brand and/or configuration during periods were sensor saturation may be occurring;
    - For low cell counts where negative values are consistently encountered, i.e. specifically for Waahi, assess whether an EXO sonde arrangement (as for river sites which have low blue-green algae cell counts but no negative fluorescence results) improves the reliability of data;
- Adding new sites of interest, especially lake sites with known issues of blue-green algae blooms;
- Including biovolume data. There is significant variability of blue-green algae sizes and biovolume will normalise these data. Furthermore, with NPS-FM guidelines based on biovolume, this will provide a more direct link between sensor fluorescence and regulatory thresholds.

7. Acknowledgements

We thank Joe Butterworth (JFB Environmental) for providing boat skipper services and assistance with sampling. We thank WRC staff (Asaeli Tulagi, Mark Hamer, Deniz Özkundakci) for provision of data and assistance with technical matters. We also thank Paula Brown (HCC) for provision of blue-green algae data for Wairoa water treatment plant intake.

8. References

Bañares-España, E., López-Rodas, V., Costas, E., Salgado, C., Flores-Moya, A., 2007. Genetic variability associated with photosynthetic pigment concentration, and photochemical and nonphotochemical quenching, in strains of the cyanobacterium Microcystis aeruginosa. FEMS Microbiol. Ecol. 60, 449–455. https://doi.org/10.1111/j.1574-6941.2007.00292.x

Brient, L., Lengronne, M., Bertrand, E., Rolland, D., Sipel, A., Steinmann, D., Baudin, I., Legeas, M., Le Rouzic, B., Bormans, M., 2008. A phycocyanin probe as a tool for monitoring cyanobacteria in freshwater bodies. J. Environ. Monit. 10, 248–255.

Gregor, J., Maršálek, B., Šípková, H., 2007. Detection and estimation of potentially toxic cyanobacteria in raw water at the drinking water treatment plant by in vivo fluorescence method. Water Res. 41, 228–234. https://doi.org/10.1016/j.watres.2006.08.011

Hodges, C.M., Wood, S.A., Puddick, J., McBride, C.G., Hamilton, D.P., 2018. Sensor manufacturer, temperature, and cyanobacteria morphology affect phycocyanin fluorescence measurements. Environ. Sci. Pollut. Res. 25, 1079–1088. https://doi.org/10.1007/s11356-017-0473-5

Izydorczyk, K., Tarczynska, M., Jurczak, T., Mrowczynski, J., Zalewski, M., 2005. Measurement of phycocyanin fluorescenceas an online early warning system for cyanobacteria in reservoir intake water. Environ. Toxicol. 20, 425–430.

Kong, Y., Lou, I., Zhang, Y., Lou, C.U., Mok, K.M., 2014. Using an online phycocyanin fluorescence probe for rapid monitoring of cyanobacteria in Macau freshwater reservoir. Hydrobiologia 741, 33–49. https://doi.org/10.1007/s10750-013-1759-3

McQuaid, N., Zamyadi, A., Prevost, M., Bird, D.F., Dorner, S., 2011. Use of in vivo phycocyanin fluorescence to monitor potential microcystin-producing cyanobacterial biovolume in a drinking water source. J. Environ. Monit. 13, 455–463. https://doi.org/10.1039/C0EM00163E

Ministry for the Environment, 2017. National Policy Statement for Freshwater Management 2014. Updated August 2017 to incorporate amendments from the National Policy Statement for Freshwater Amendment Order 2017.

Ministry of Health, 2018. Guidelines for Drinking-water Quality Management for New Zealand (6th edition June 2018). Wellington.

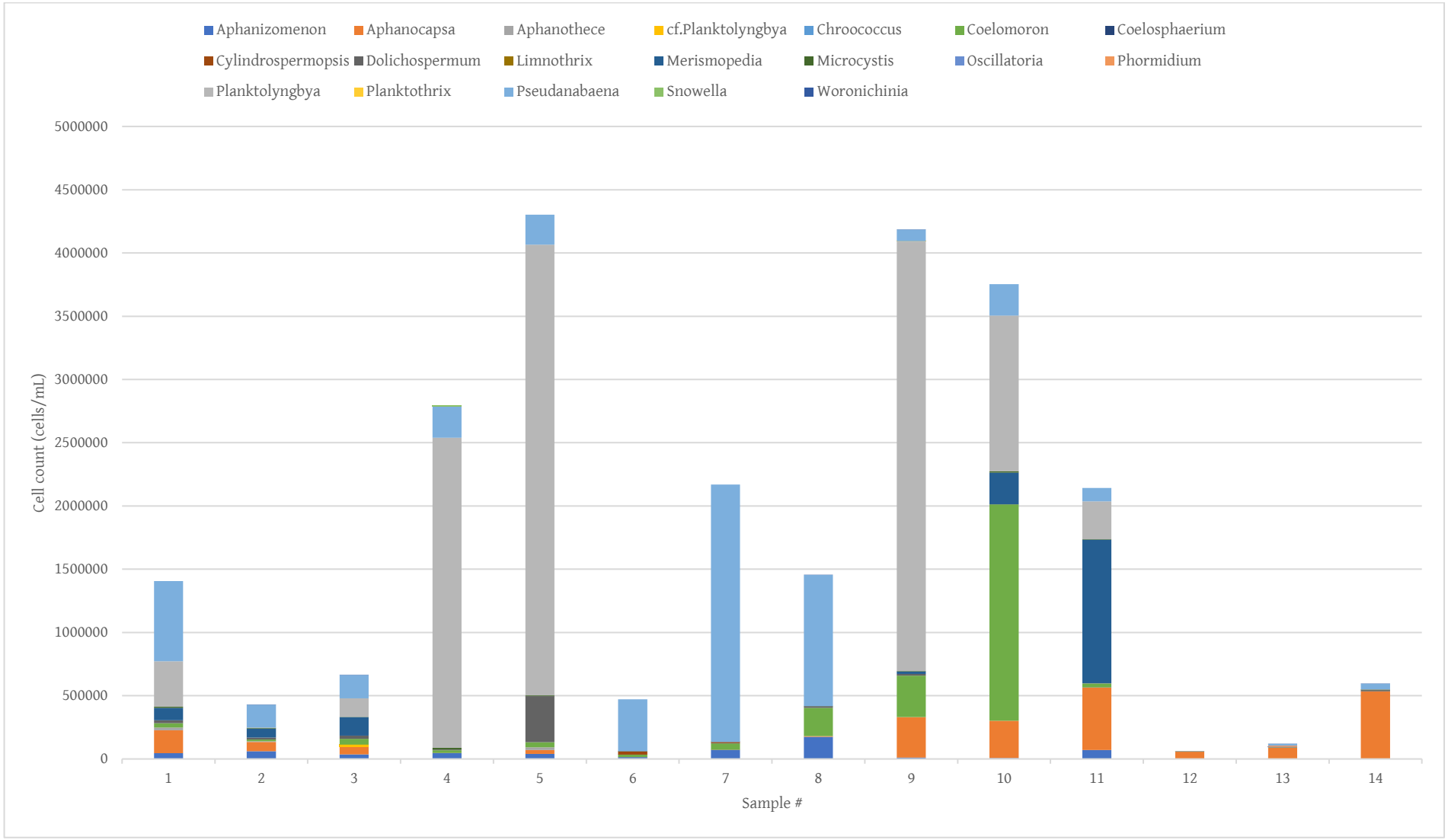
Zamyadi, A., 2011. The value of in vivo monitoring and chlorination for the control of toxic cyanobacteria in drinking water production. PhD thesis, Ecole polytechnique de Montreal—Universite de Montreal/University of Montreal.

Zamyadi, A., McQuaid, N., Dorner, S., Bird, D.F., Burch, M., Baker, P., Hobson, P., Prevost, M., 2012a. Cyanobacterial detection using in vivo fluorescence probes: Managing interferences for improved decision-making. J. - Am. Water Works Assoc. 104, E466–E479. https://doi.org/10.5942/jawwa.2012.104.0114

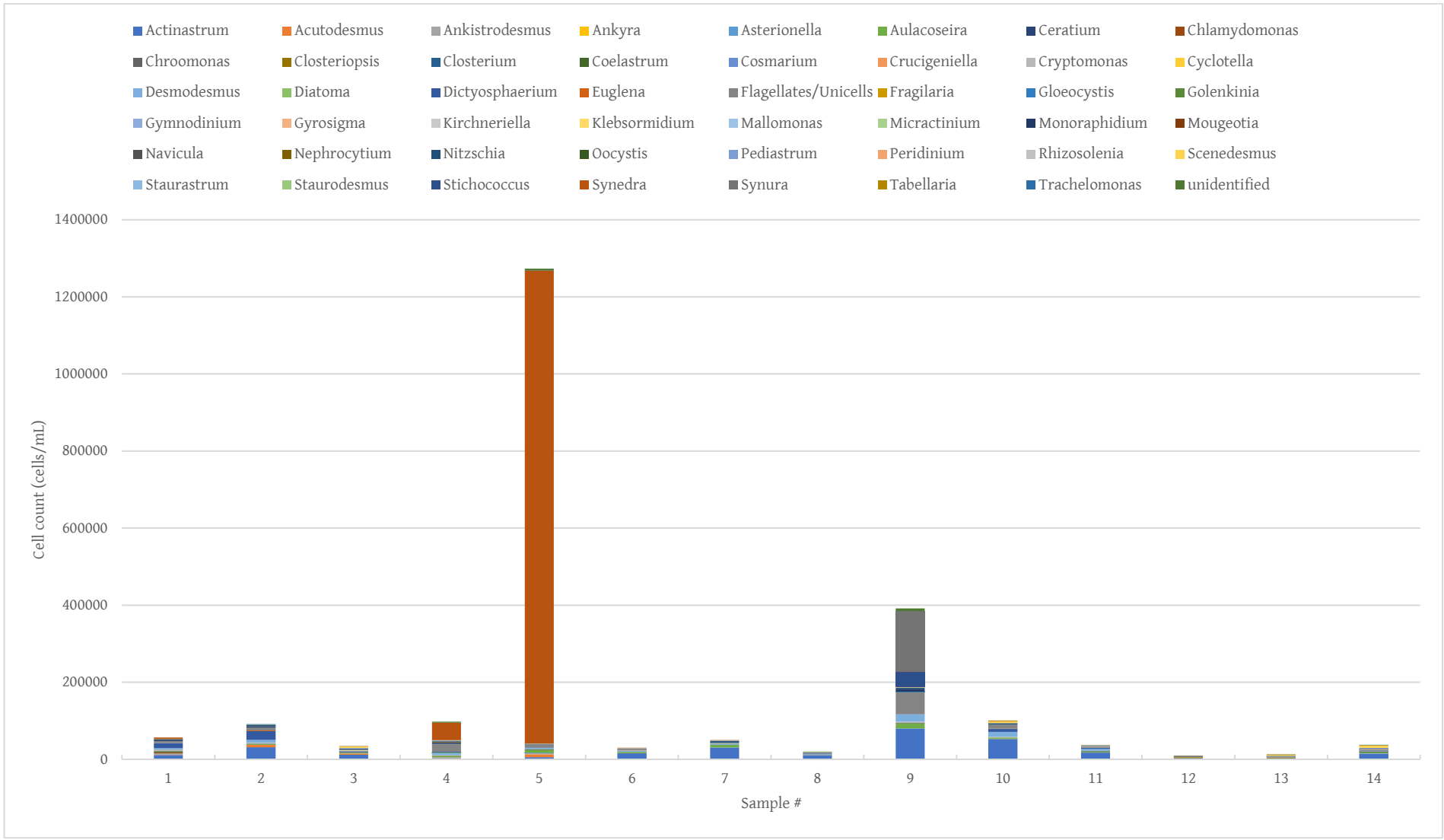
Zamyadi, A., McQuaid, N., Prevost, M., Dorner, S., 2012b. Monitoring of potentially toxic cyanobacteria using an online multi-probe in drinking water sources. J. Environ. Monit. 14, 579–588. https://doi.org/10.1039/C1EM10819K

Appendix 1 Pivot Graphs

Blue-green algae - genus

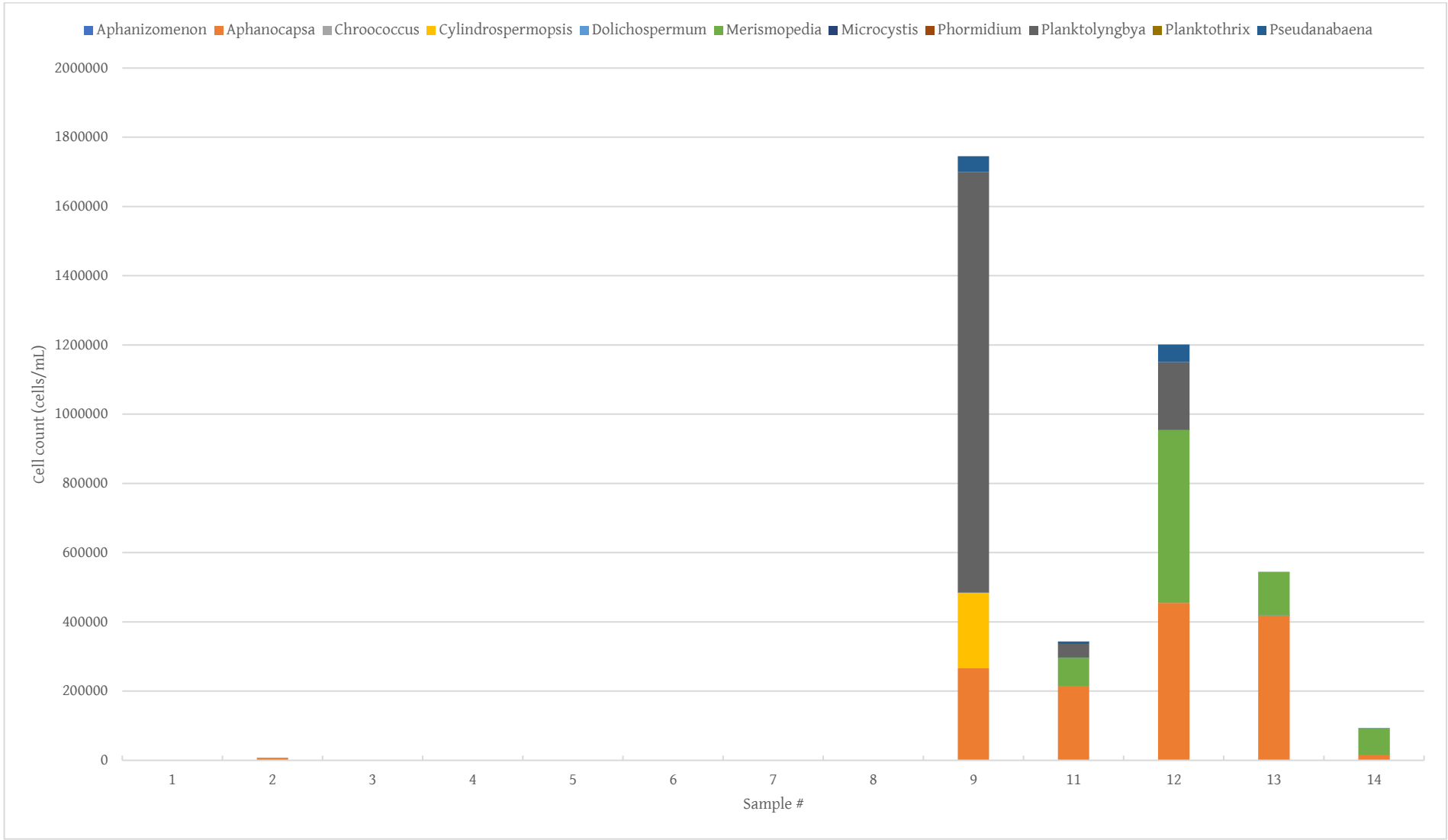


“Non” blue-green algae - genus

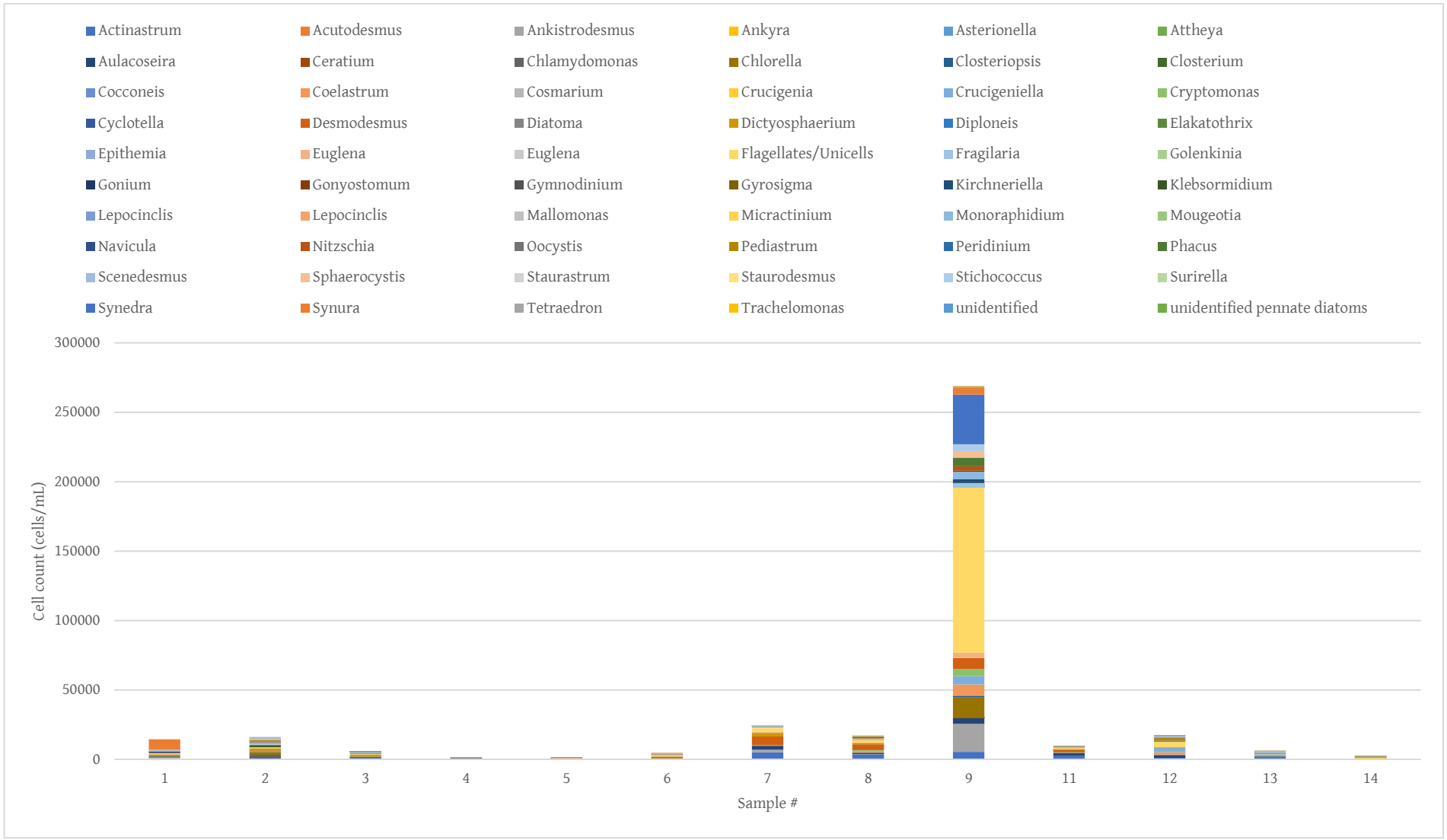


Whangape

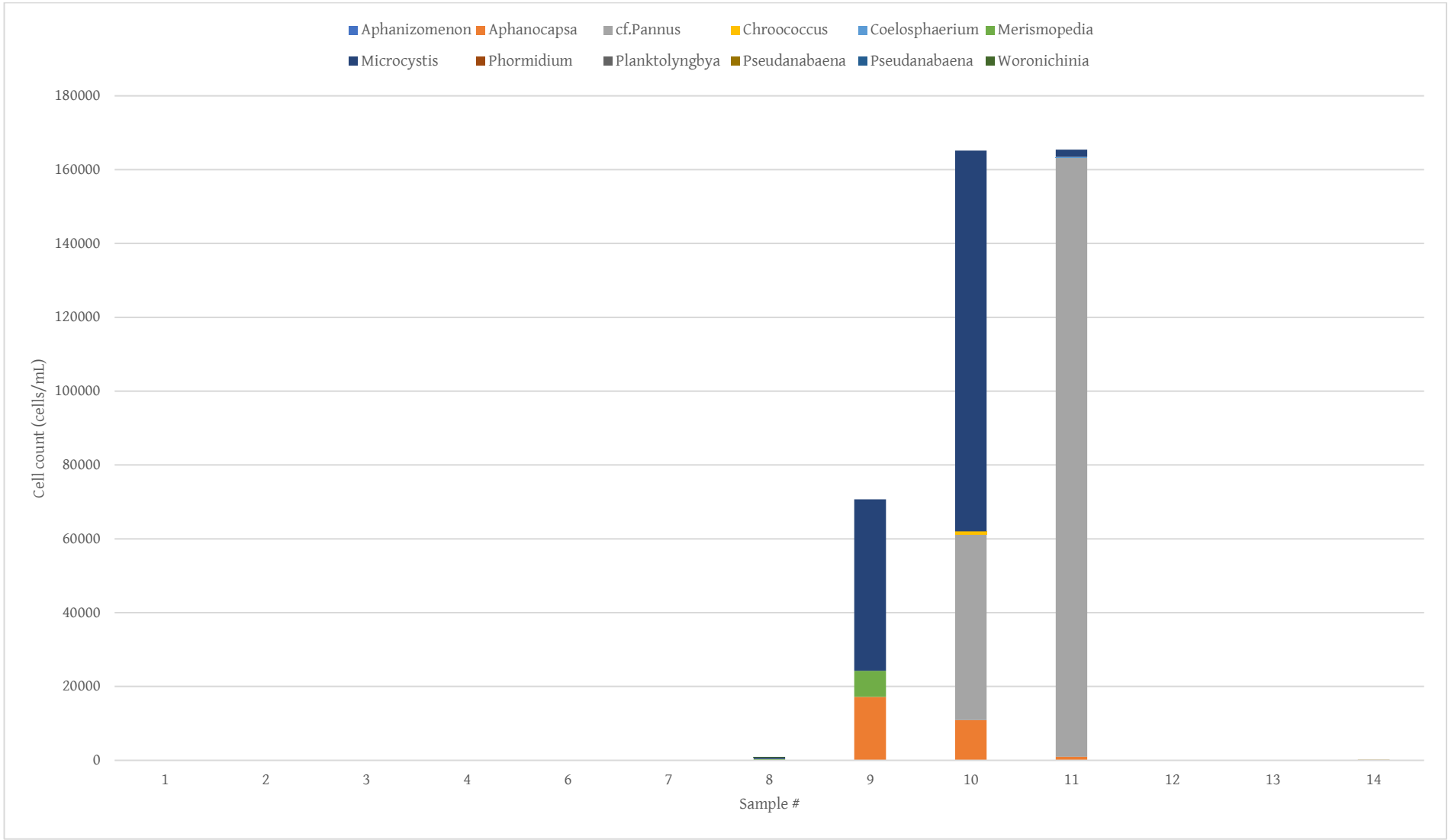
Blue-green algae – genus



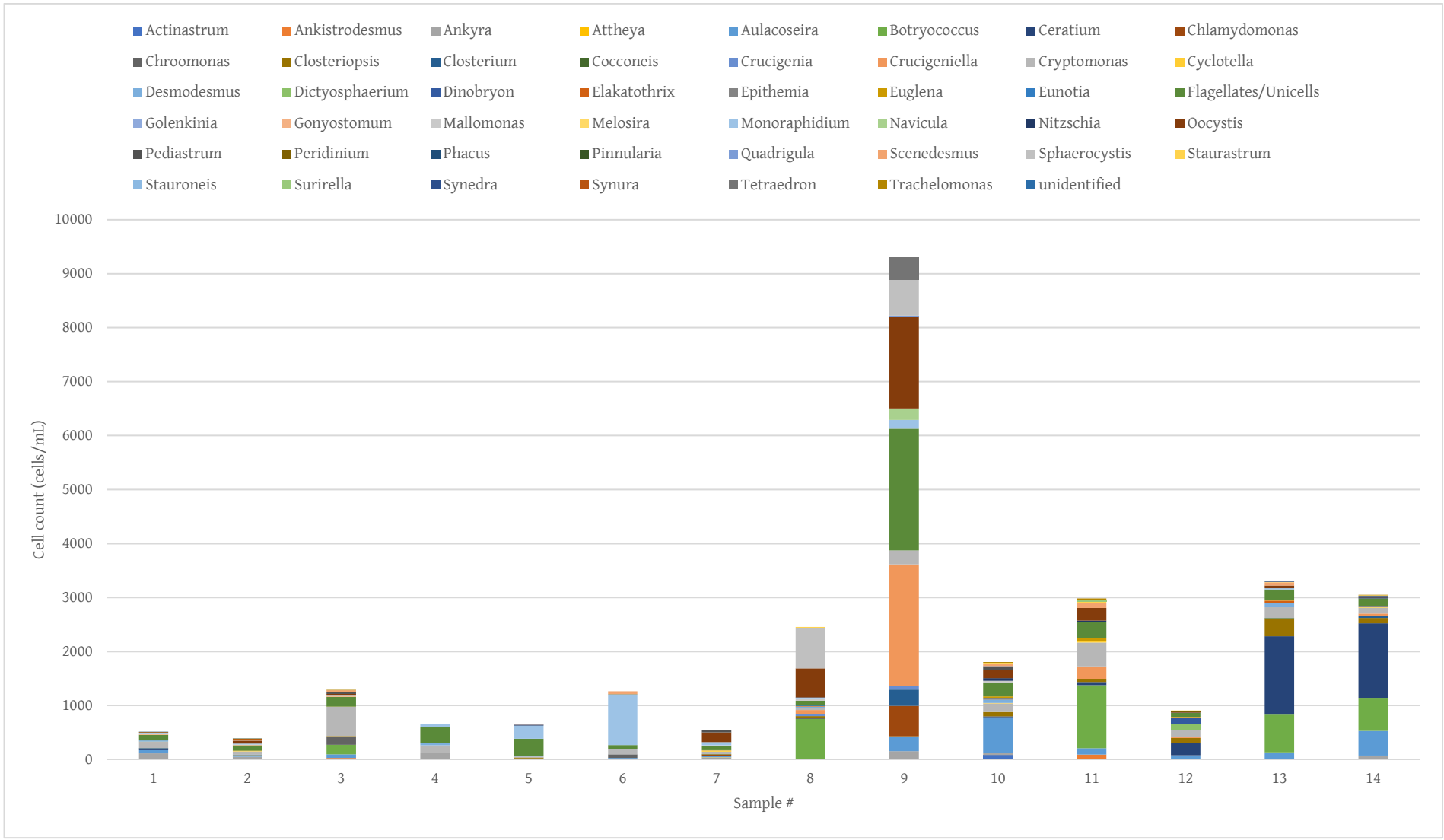
“Non” blue-green algae – genus



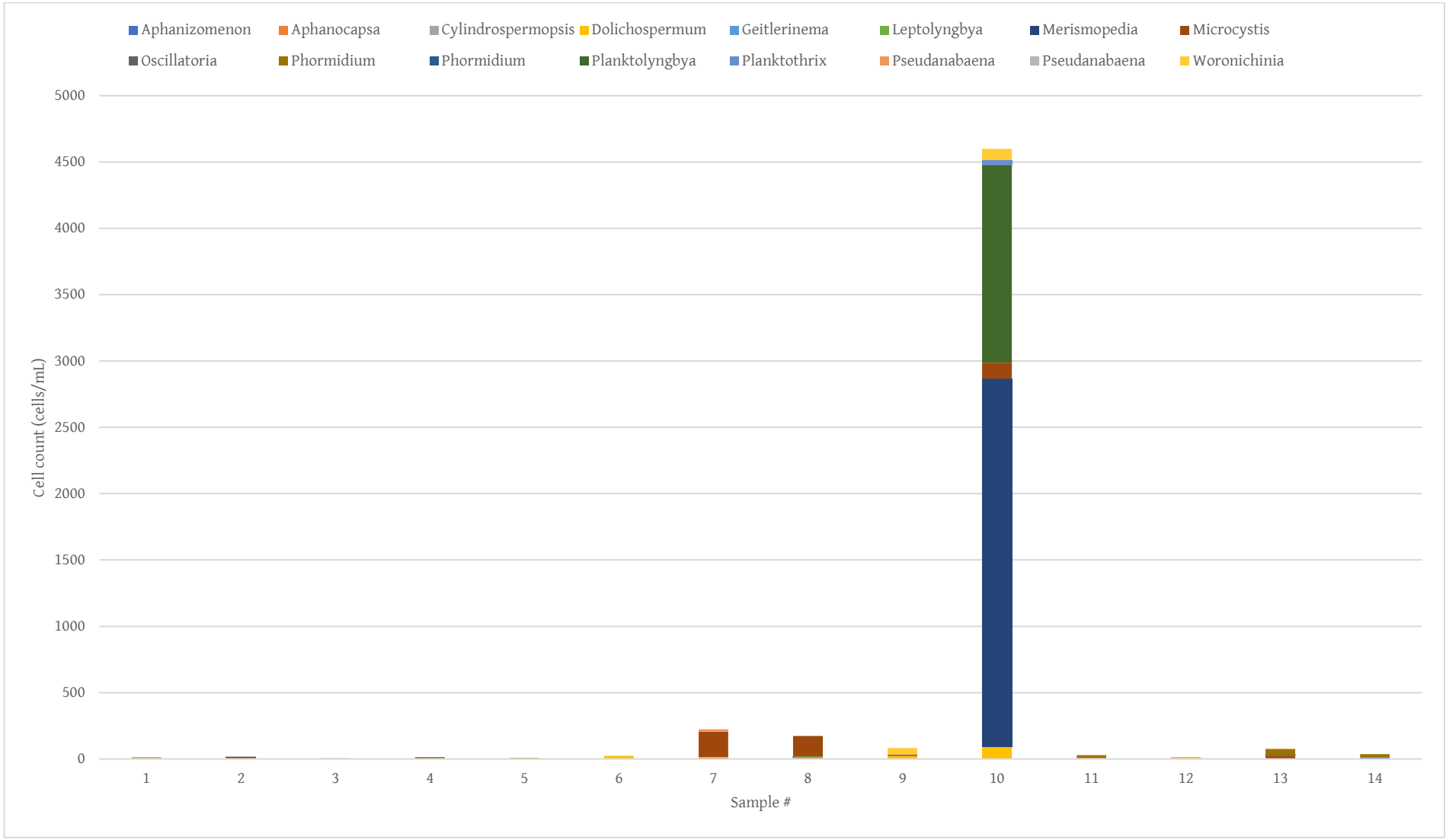
Blue-green algae – genus



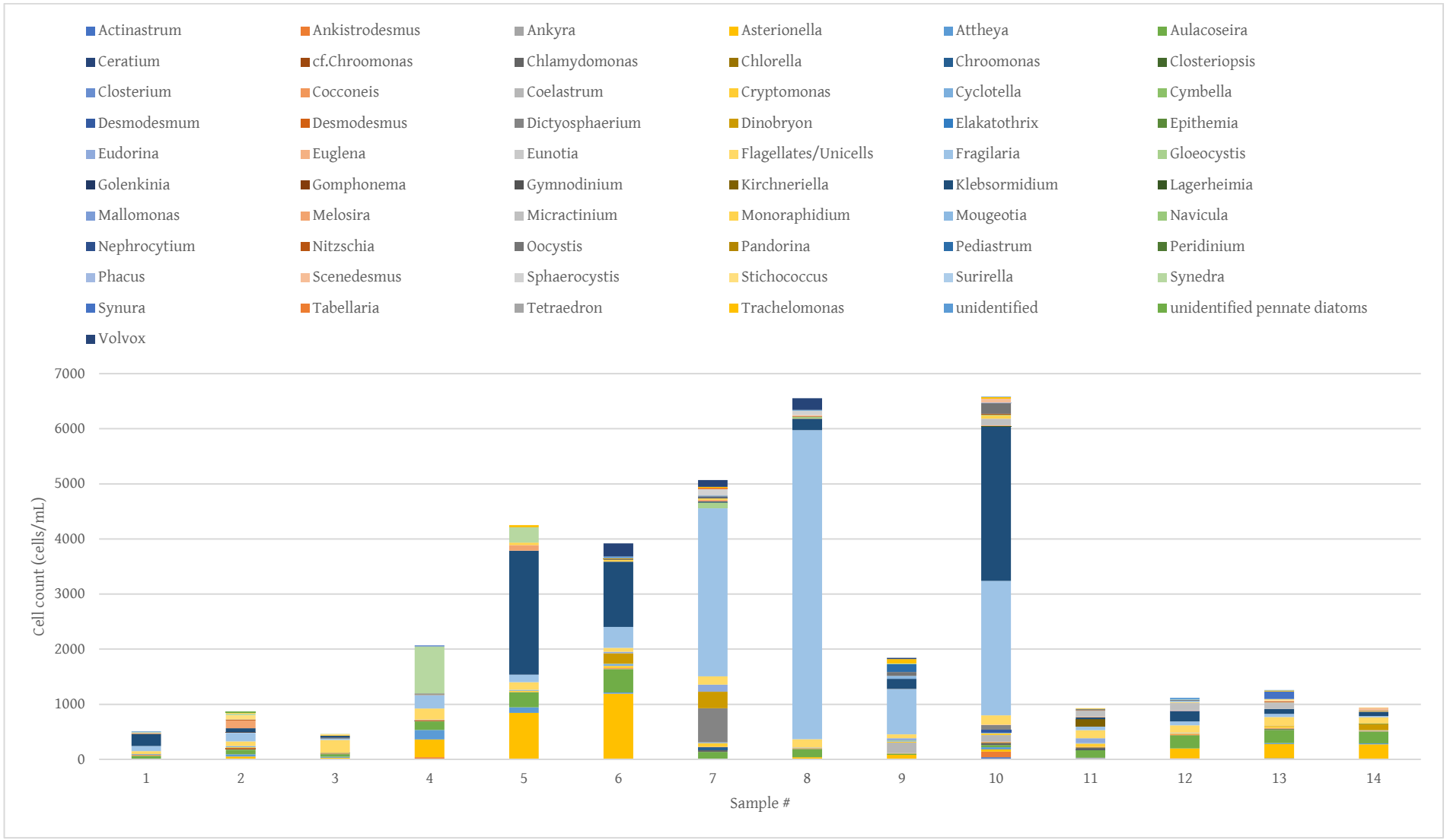
“Non” blue-green algae – genus



Blue-green algae – genus

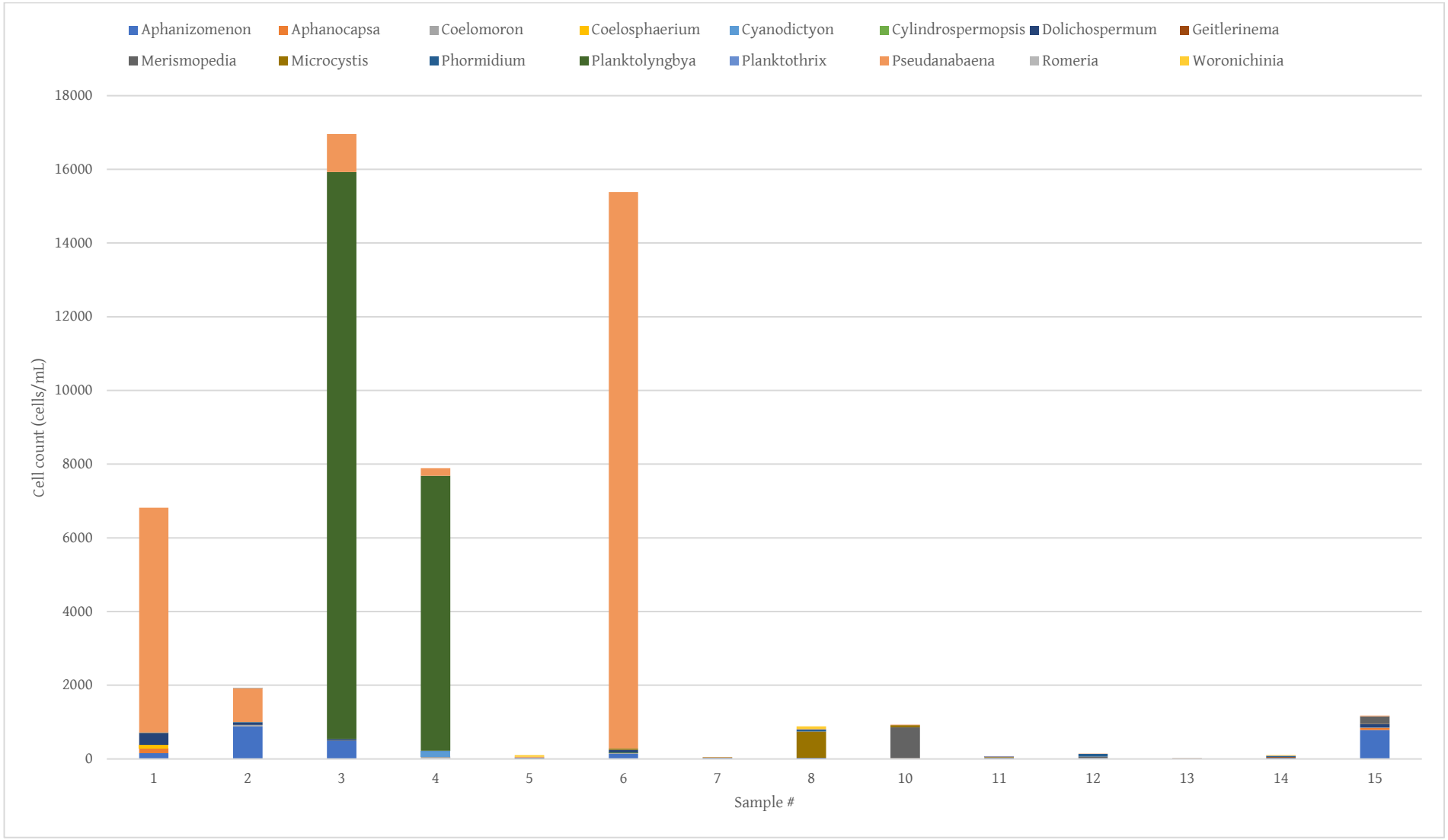


“Non” blue-green algae – genus

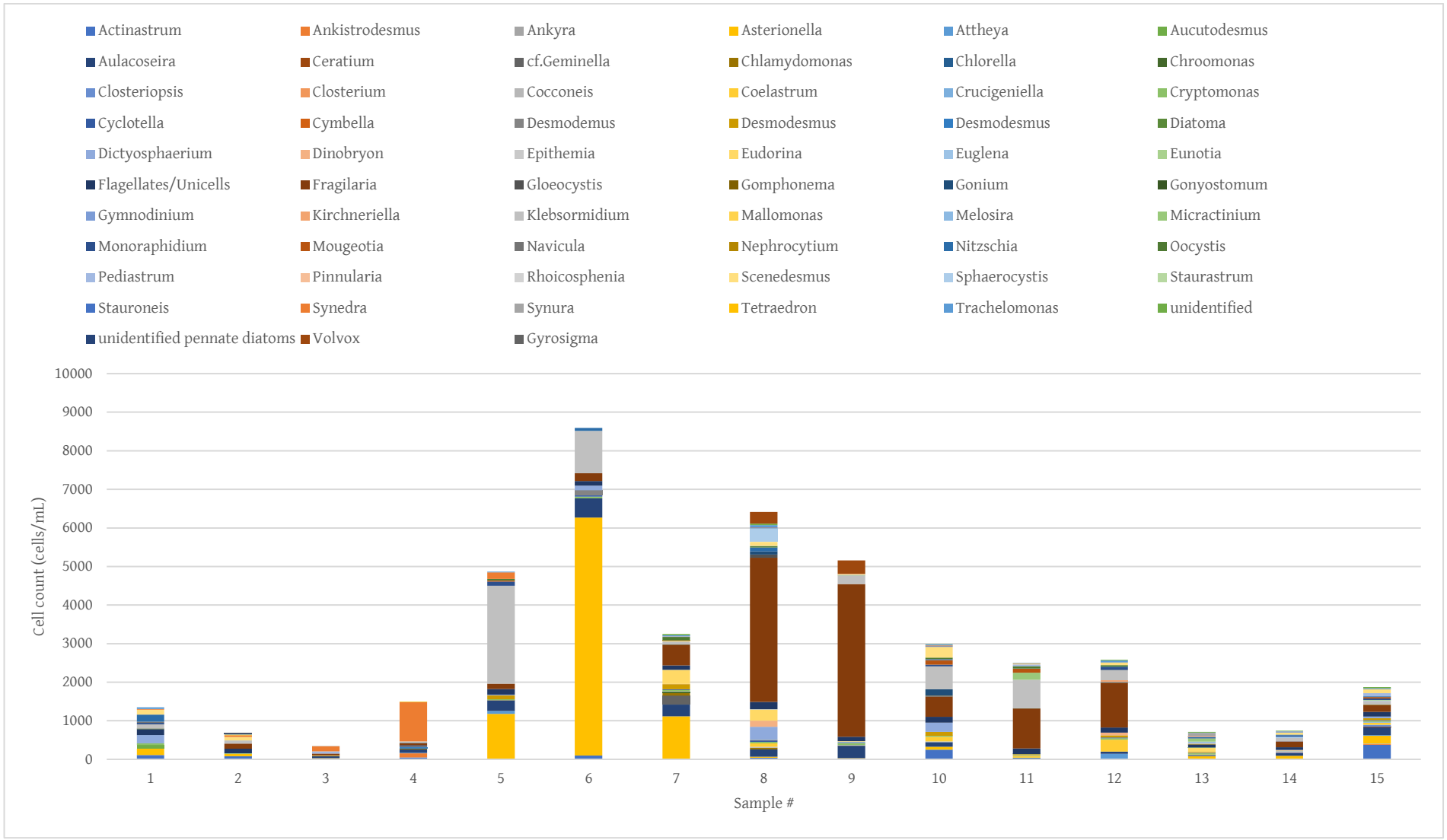


Elbow

Blue-green algae – genus

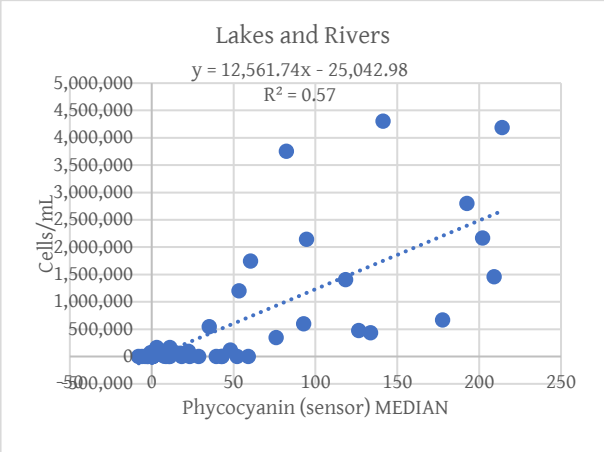


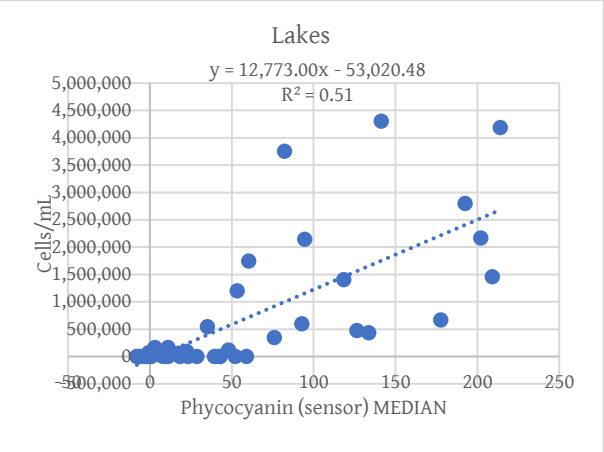
“Non” blue-green algae – genus



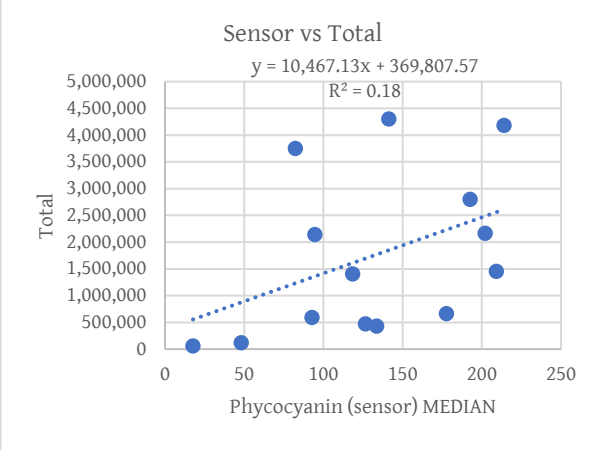
Appendix 2 Regression graphs for blue-green algae cell counts vs PC sensor fluorescence

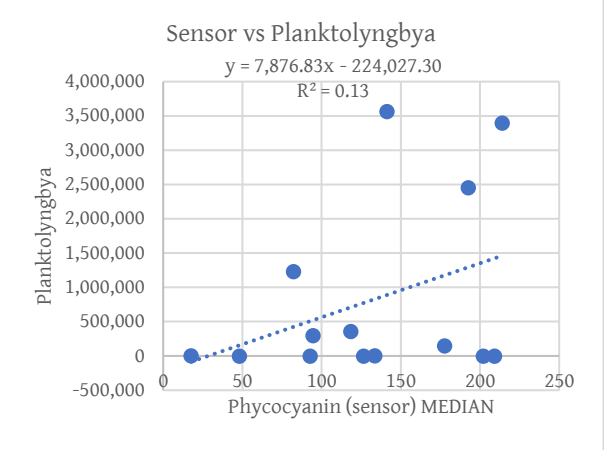
Lakes and Rivers combined – Total (cells/mL)

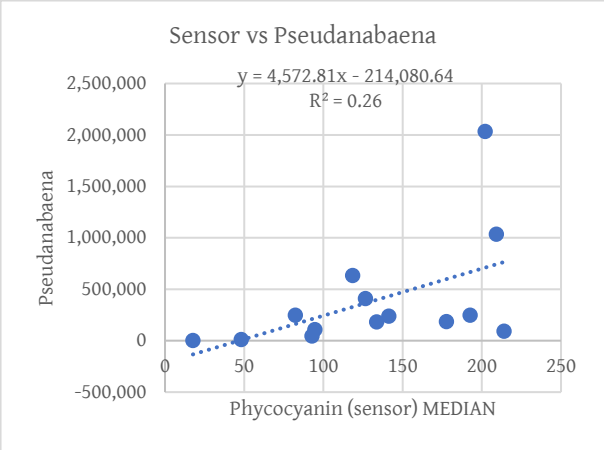


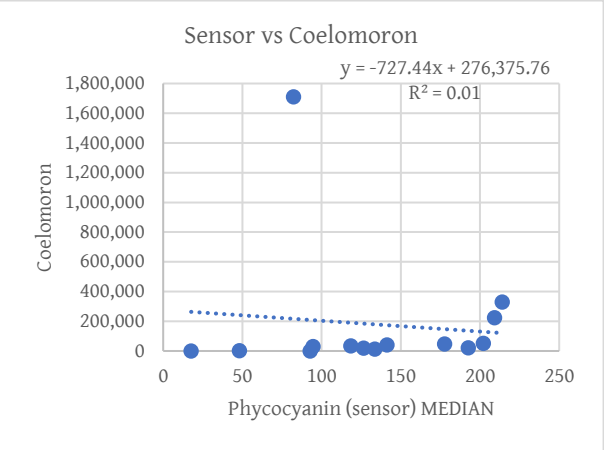


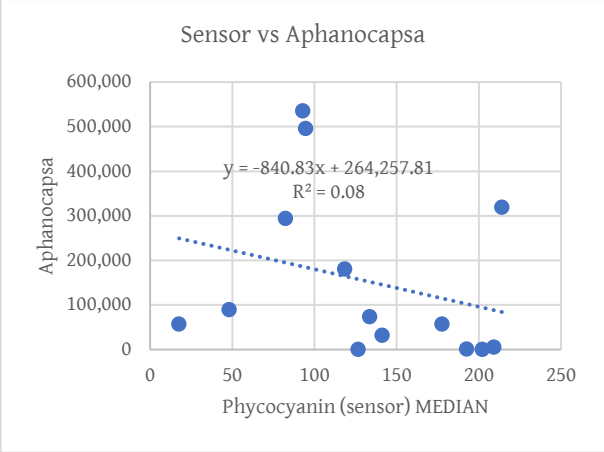
Waikare – Total, dominant genus and species (cells/mL)

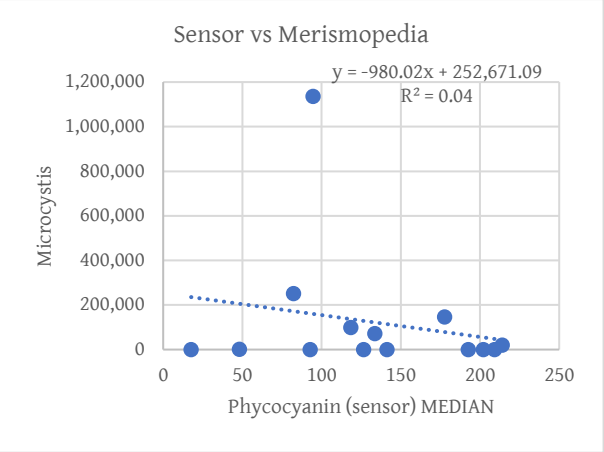


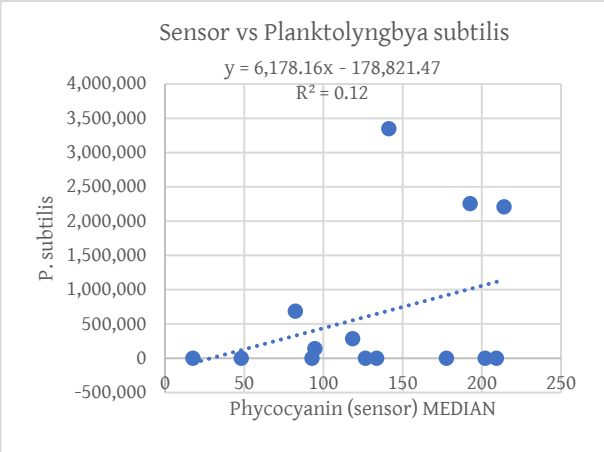


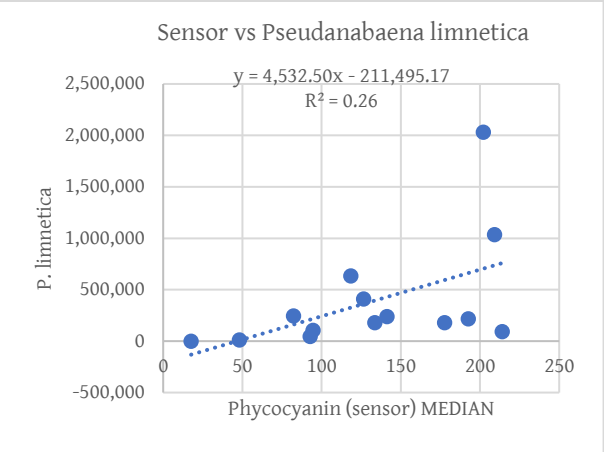


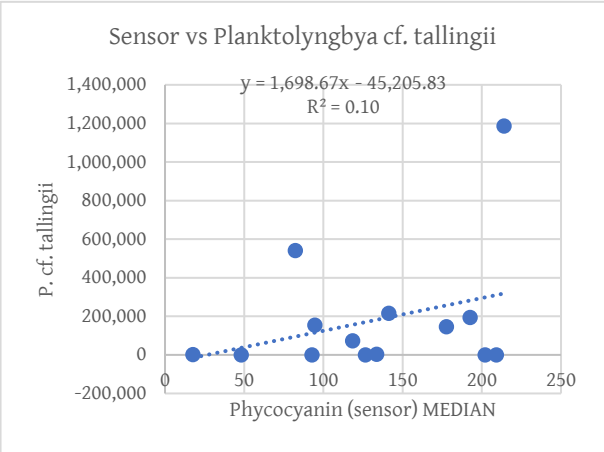


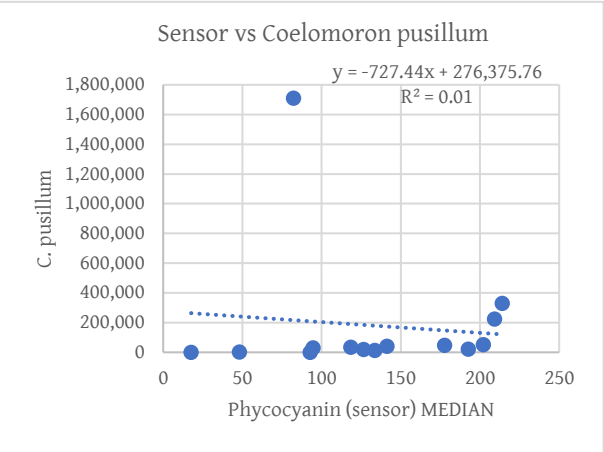

Sensor vs Aphanocapsa
y = -840.83x + 264,257.81
R² = 0.08
Aphanocapsa
Phycocyanin (sensor) MEDIAN


Sensor vs Merismopedia
y = -980.02x + 252,671.09
R² = 0.04
Microcystis
Phycocyanin (sensor) MEDIAN

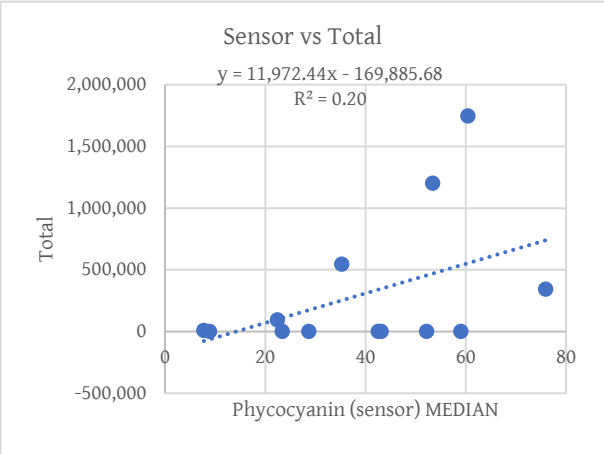

Sensor vs Planktolyngbya subtilis
y = 6,178.16x - 178,821.47
R² = 0.12
P. subtilis
Phycocyanin (sensor) MEDIAN

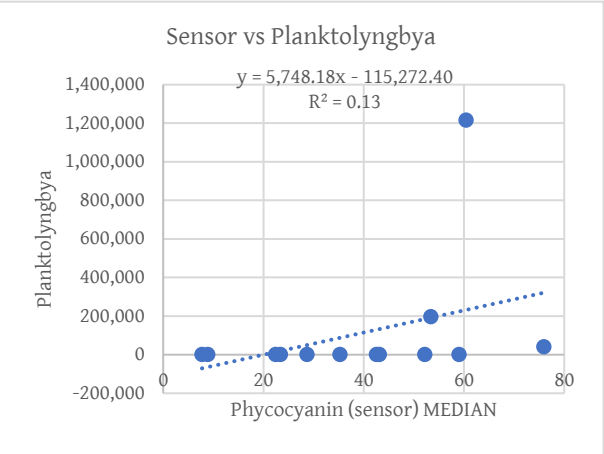

Sensor vs Pseudanabaena limnetica
y = 4,532.50x - 211,495.17
R² = 0.26
P. limnetica
Phycocyanin (sensor) MEDIAN

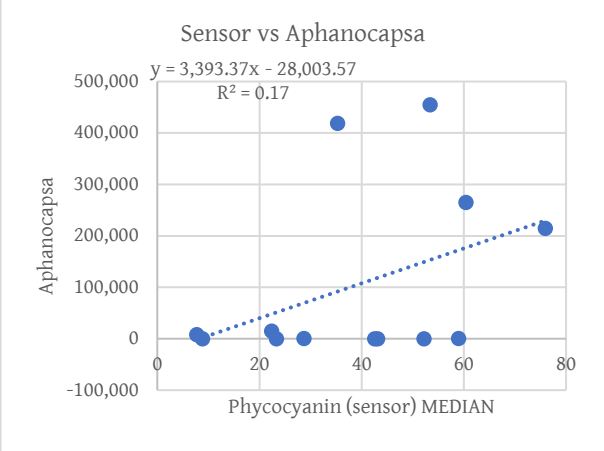

Sensor vs Planktolyngbya cf. tallingii
y = 1,698.67x - 45,205.83
R² = 0.10
P. cf. tallingii
Phycocyanin (sensor) MEDIAN

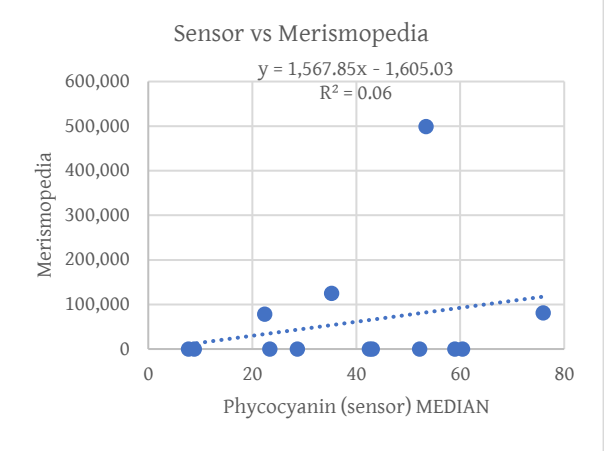

Sensor vs Coelomoron pusillum
y = -727.44x + 276,375.76
R² = 0.01
C. pusillum
Phycocyanin (sensor) MEDIAN

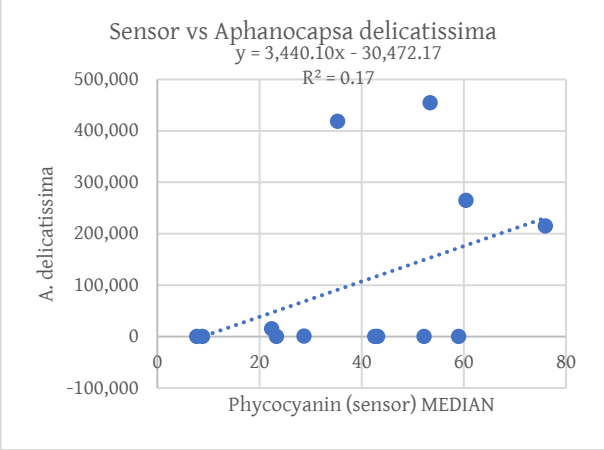
Whangape – Total, dominant genus and species (cells/mL)

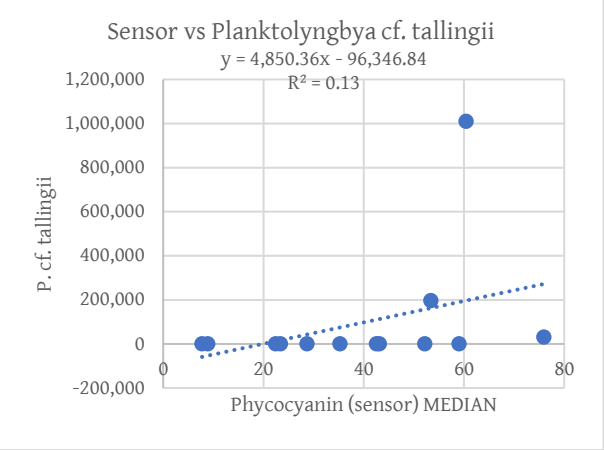




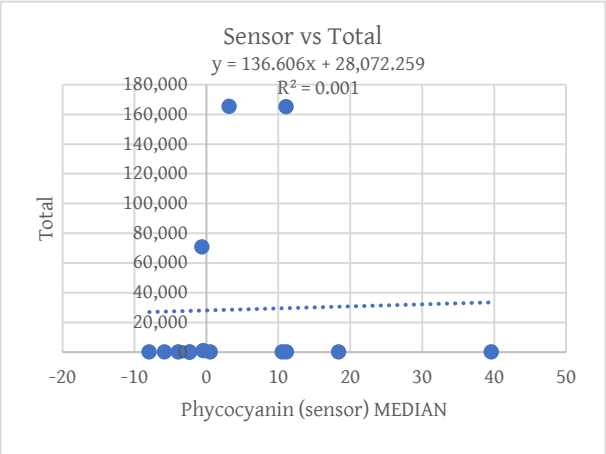


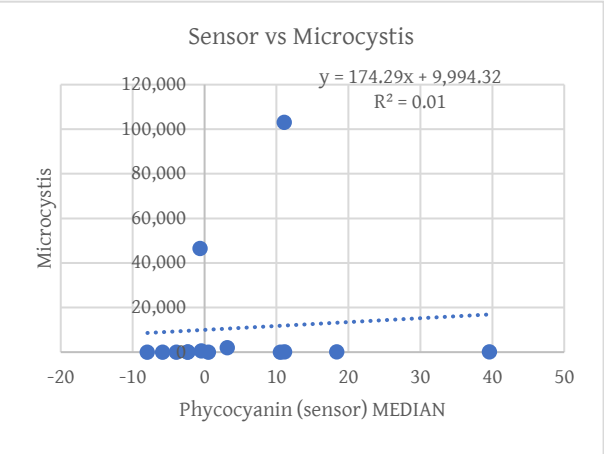


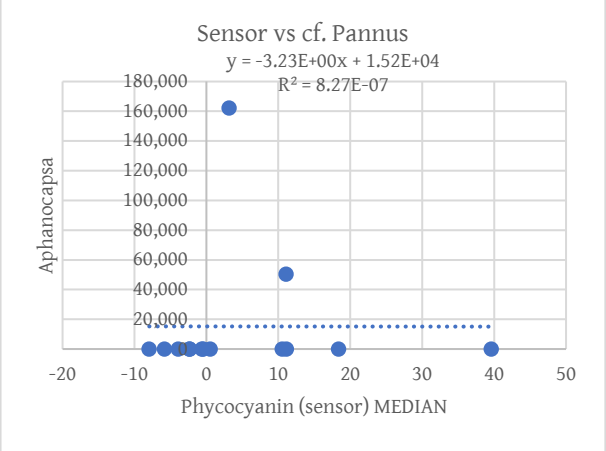


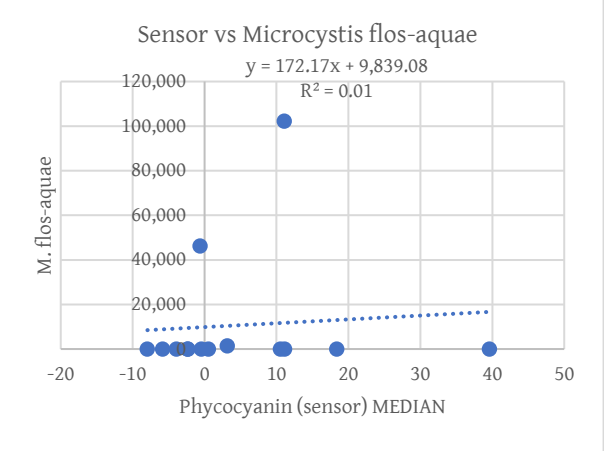


Waahi – Total, dominant genus and species (cells/mL)

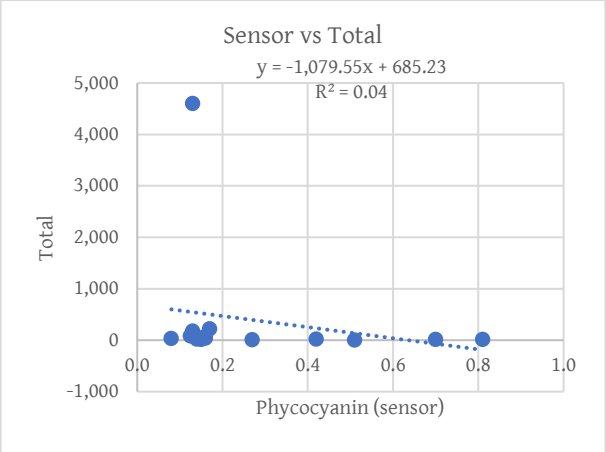


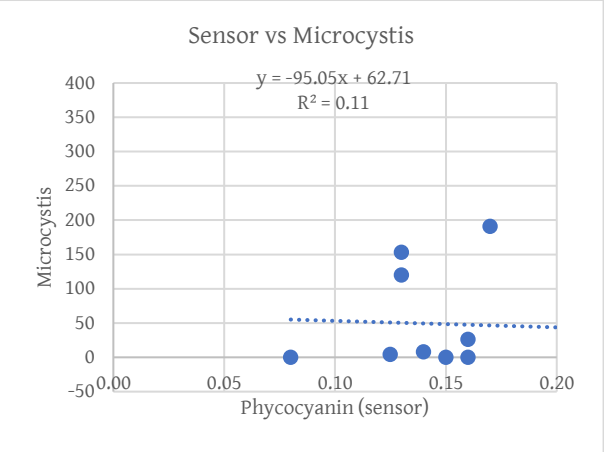






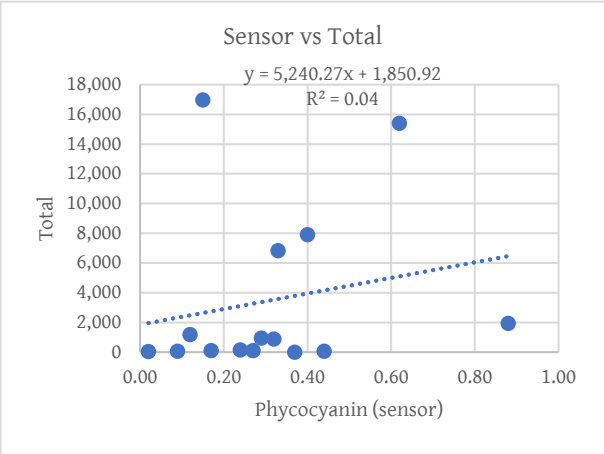
Hamilton Traffic – Total and dominant genus (cells/mL)[17]

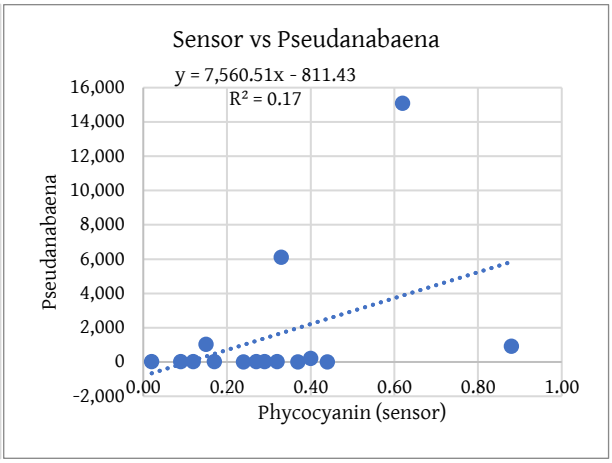




[17] Merismopedia and Planktolyngbya are not included as they were present for only 1 and 2 months respectively. Regressions were not undertaken to species level due to weak regressions at genus level.

Elbow – Total and dominant genus (cells/mL)[18]

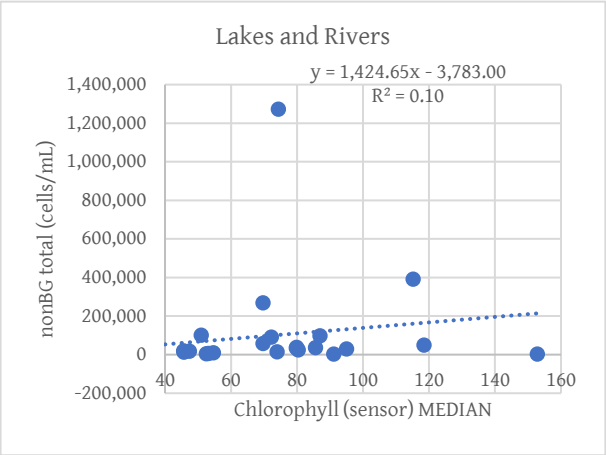


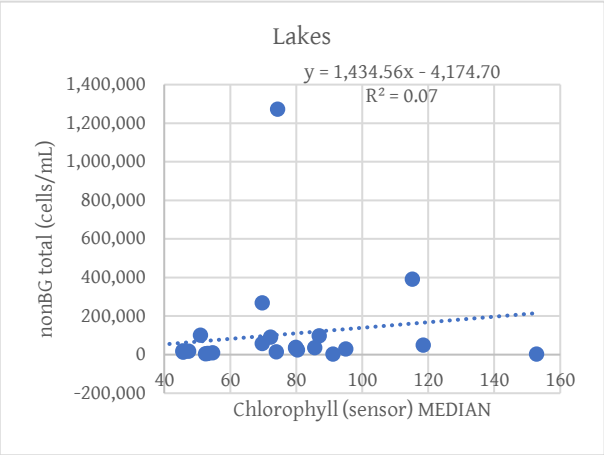


[18] Planktolyngbya was not included as it was present for only 3 months. Regressions were not undertaken to species level as Pseudanabaena was almost exclusively *Pseudanabaena limnetica*.

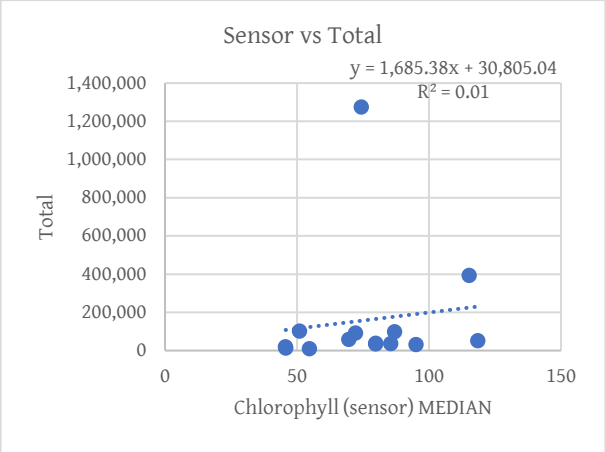
Appendix 3 Regression graphs for "non" blue-green algae cell counts vs Chl-*a* sensor fluorescence

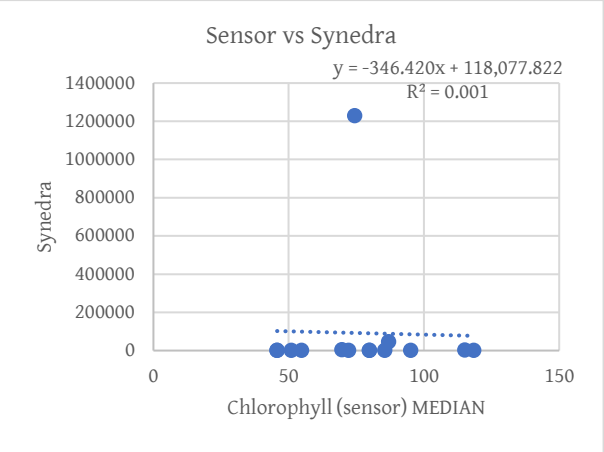
Lakes and Rivers combined – Total (cells/mL)

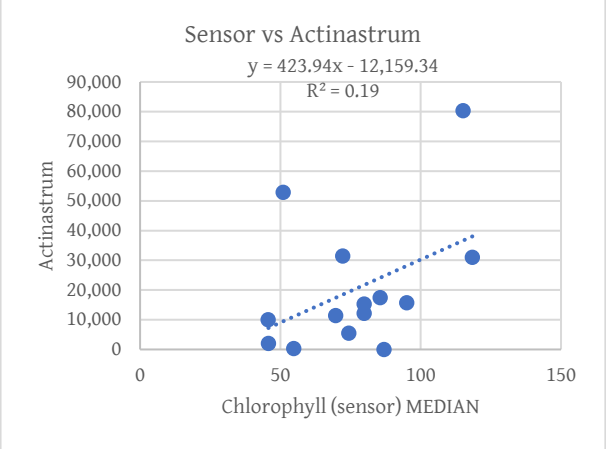




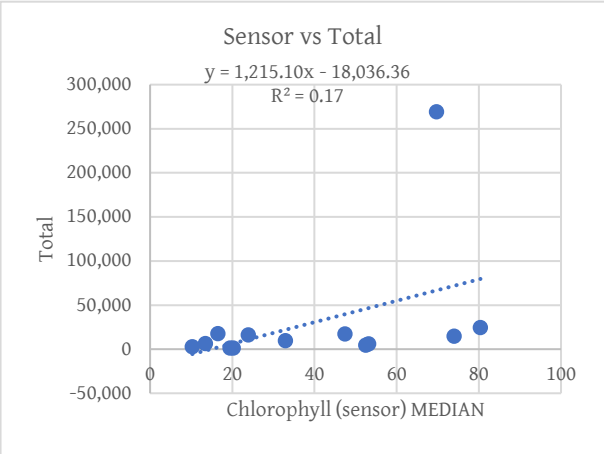
Waikare – Total and dominant genus (cells/mL)

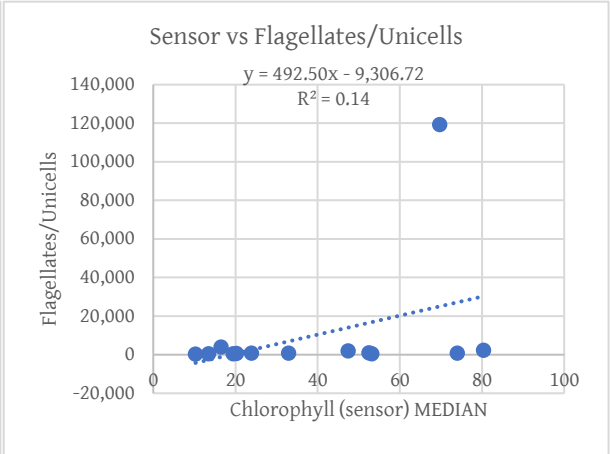


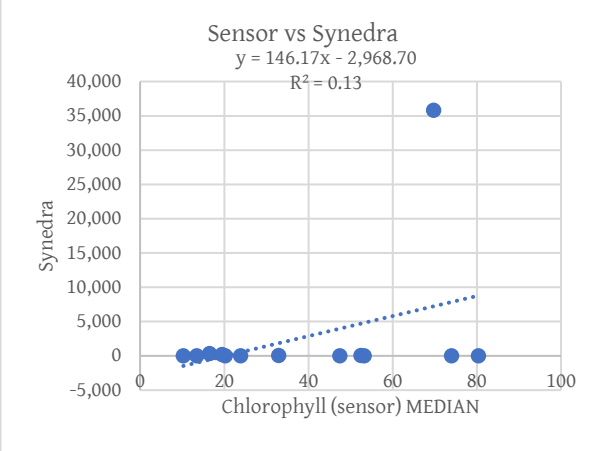


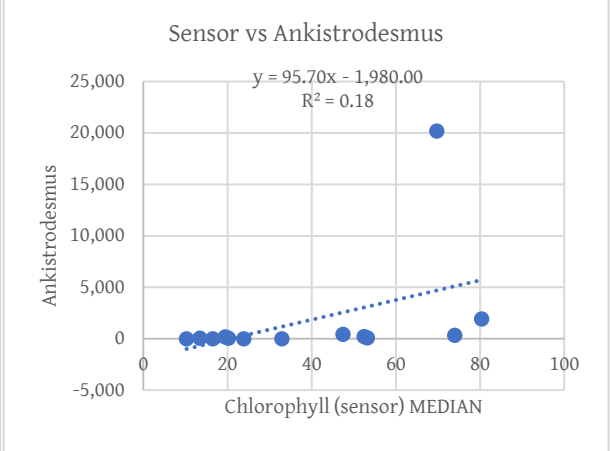


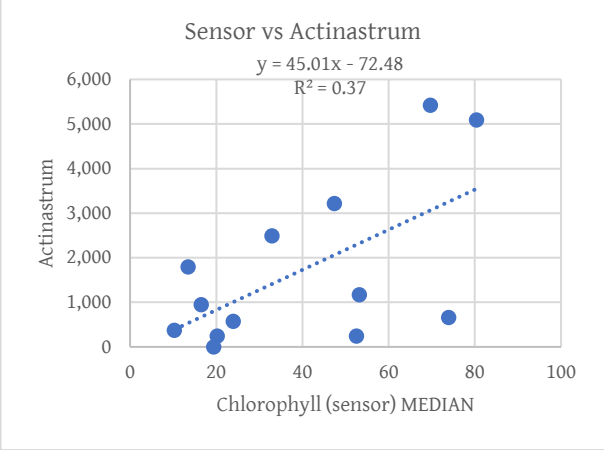
Whangape – Total and dominant genus (cells/mL)



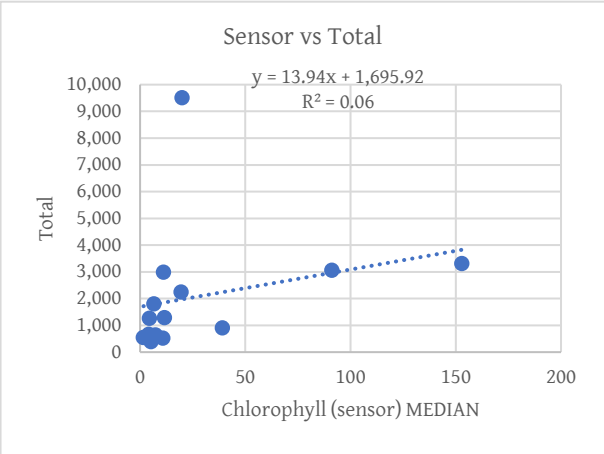


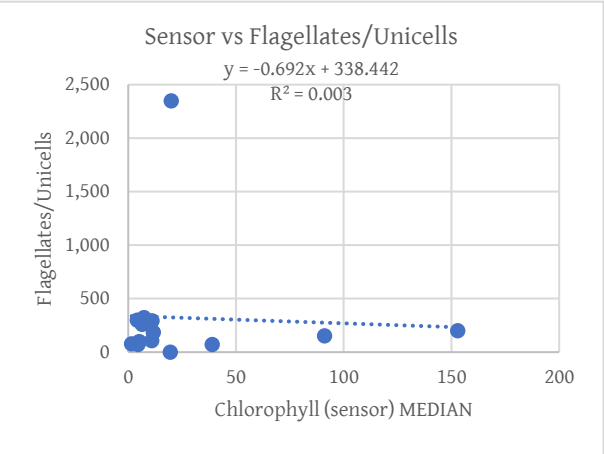


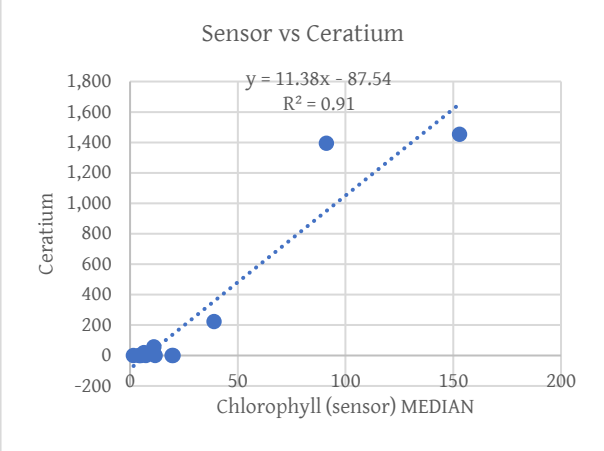


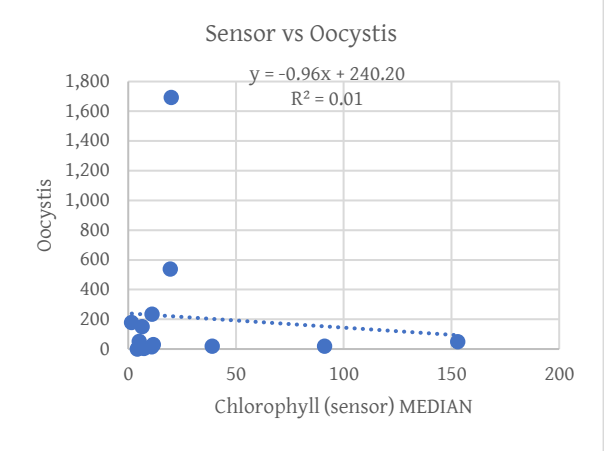


Waahi – Total and dominant genus (cells/mL)

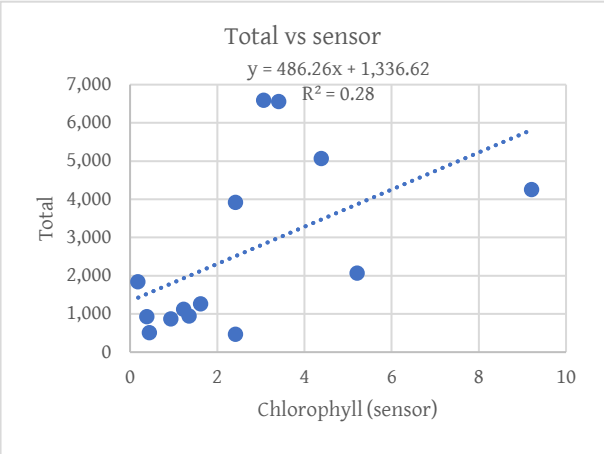


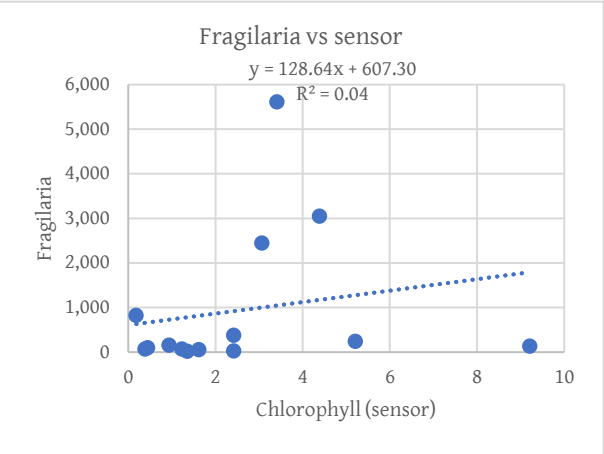


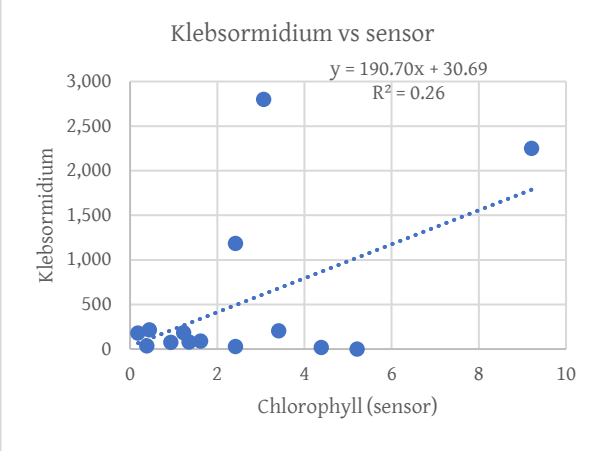


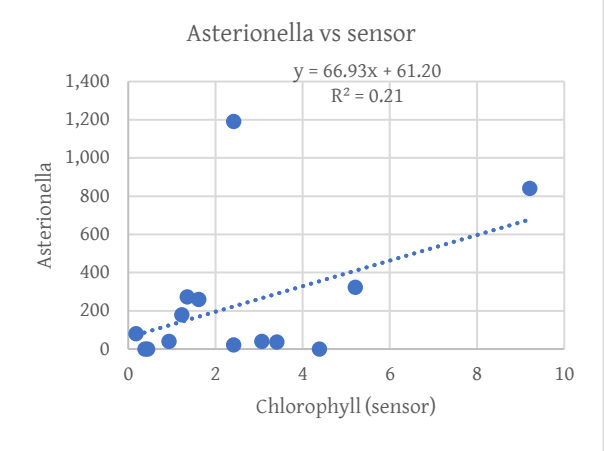


Hamilton Traffic – Total and dominant genus (cells/mL)

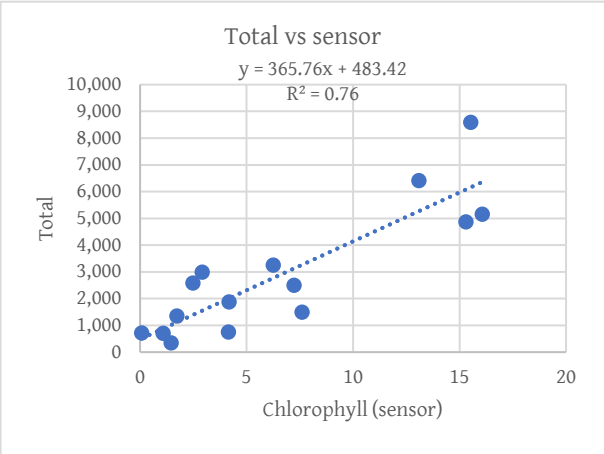


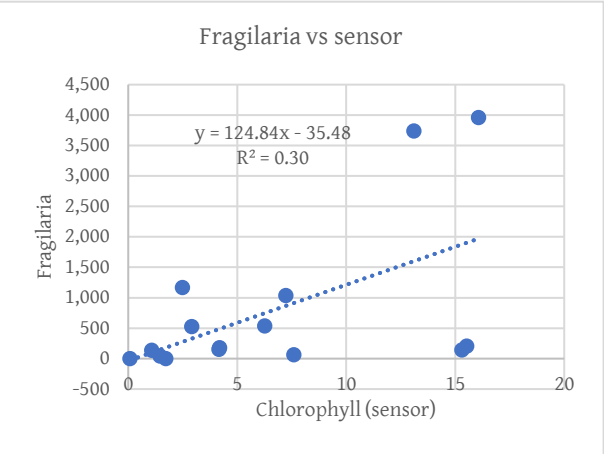


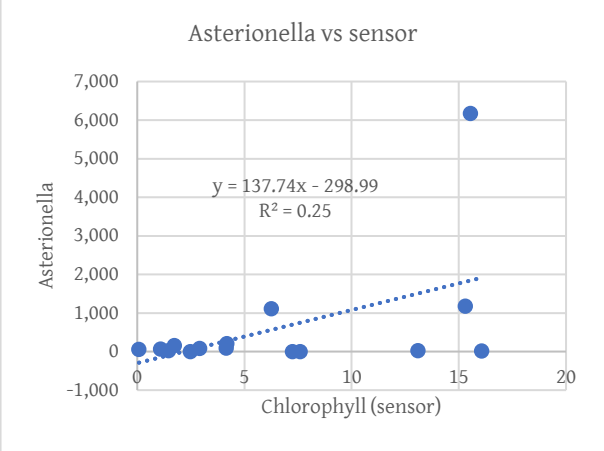


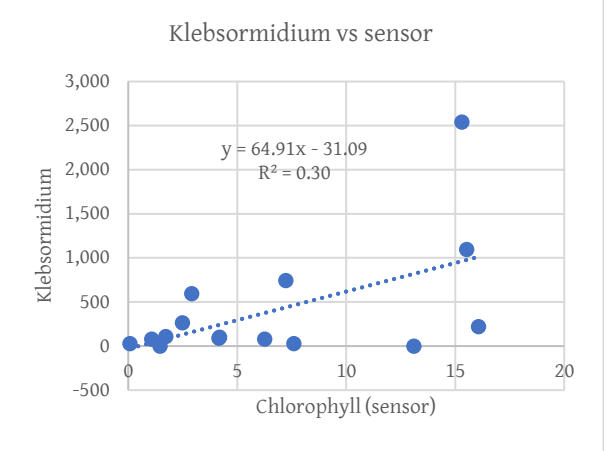


Elbow – Total and dominant genus (cells/mL)









Appendix 4 Regression graphs for turbidity (sensor) vs turbidity (WQ meter)